

UMA INTRODUÇÃO À EXPERIMENTAÇÃO FLORESTAL

WALDENEI TRAVASSOS DE QUEIROZ





WALDENEI TRAVASSOS DE QUEIROZ

UMA INTRODUÇÃO À EXPERIMENTAÇÃO FLORESTAL

Belém
2023

©2023 Universidade Federal Rural da Amazônia

MINISTÉRIO DA EDUCAÇÃO

Ministro: Camilo Sobreira de Santana

UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA

Reitora: Herdjanira Veras de Lima

Vice-Reitor: Jaime Viana de Sousa

GERENTE DE EDITORAÇÃO

Suely Nazaré Furtado França

COMISSÃO EDITORIAL

Antônia Benedita da Silva Bronze

Israel Hidenburgo Aniceto Cintra

Leony Soares Marinho

Suely Nazaré Furtado França

Tabila Verena da Silva Leite

EQUIPE EDITORIAL

Cristiane do Espírito Santo Coelho

Socorro de Fátima Souza da Silva Viegas

Inácia Faro Libonati

Felipe do Nascimento Barata

REVISÃO TEXTUAL E ORTOGRÁFICA

Cleide Lúcia Gaspar da Assunção

DIAGRAMAÇÃO

Osimar R. Araujo (osiarte@gmail.com)

Todos os direitos reservados

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.160/98)



Av. Tancredo Neves, 2501, CEP: 66077-830 - Terra Firme

E- mail: editora@ufra.edu.br

Editora Associada à Associação Brasileira das Editoras Universitárias – ABEU

Dados Internacionais de Catalogação na Publicação (CIP)

Queiroz, Waldenei Travassos de.

Uma introdução à experimentação florestal / Waldenei

Travassos de Queiroz.- Belém: Universidade Federal Rural
da Amazônia, 2023.

442. p.: il.

ISBN: 978-65-994846-9-8

DOI: <https://doi.org/10.4322/978-65-994846-9-8>

1. Experimentação - Florestal. 2. Estatística -
Matemática. 3. Engenharia Florestal. I. Título.

CDD 23. ed. 634.928

Elaborado por: Cristiane do Espírito Santo Coelho - CRB2/1027

À minha amada esposa Izabel;
Aos meus filhos Anderson e Alex;
Às minhas noras Alessandra e Lorena;
Aos meus netos Lucas, Gabriel e Alicia;
Aos meus pais Wagner e Erecina *in memoriam*



AGRADECIMENTOS

Agradeço à bibliotecária Suely Nazaré Furtado França, gerente de editoração da EDUFRA e à sua equipe pelo empenho e dedicação na publicação deste livro e à Professora Cleide Lúcia Gaspar da Assunção pela cuidadosa revisão linguística e ortográfica do texto final.

Os meus sinceros agradecimentos à Universidade Federal Rural da Amazônia (UFRA), que, quando ainda Faculdade de Ciências Agrárias do Pará (FCAP), proporcionou-me o desempenho do exercício do magistério por 41 anos. Nesse longo período de tempo eu pude desenvolver atividades de extensão, pesquisa e ensino, mas, principalmente aprimorar as técnicas sobre planejamento e análise de experimentos florestais, possibilitando uma melhor formação dos estudantes, futuros Engenheiros Florestais, para que, com melhor preparo eles pudessem desenvolver suas atividades profissionais nos mais diferentes ramos de sua profissão, principalmente na Amazônia. Por outro lado, na condição de professor de alunos de graduação e pós-graduação, na UFRA e em outras Instituições de ensino e pesquisa da região amazônica, oportunidades especiais de convivência com estudantes, colegas professores e servidores foram proporcionadas a mim. Isso me fez compreender melhor as bases que embelezam as relações humanas, condição importante para deslindar os segredos da nossa missão de vida.

Ao Doutor Sylvio Péllico Netto (UFPR), meu professor e amigo de longas datas, pelos conhecimentos ensinados na área de experimentação florestal e, também, aos professores doutores da ESALQ-USP: Frederico Pimentel Gomes (*in memoriam*), Humberto de Campos (*in memoriam*), Décio Barbin e Cassio Roberto de Melo Godoi, pelos ensinamentos que deles recebi na área de estatística e experimentação, em nossa formação de pós-graduação. Quero expressar meus agradecimentos também ao Professor Doutor Carlos Tadeu dos Santos Dias pelos conhecimentos transmitidos do Sistema de Análise Estatística SAS. Agradeço ainda ao Professor Doutor Hilton Thadeu Zarate do Couto (*in memoriam*) pelo incentivo, ensinamentos e amizade permanente que começou quando da realização de meu doutorado na ESALQ-USP.

Meus agradecimentos especiais, aos revisores deste trabalho por seus conhecimentos e tempo disponibilizado para as correções e sugestões apresentadas que muito contribuíram para a melhoria da presente publicação:

À Engenheira Florestal Anabel Aparecida de Mello, Professora da UFS, Mestra pela UFLA e Doutora pela UFPR;

Ao Engenheiro Florestal Francisco José de Barros Cavalcanti, Professor da UFRRJ, Mestre pelo INPA e Doutor pela UFPR;

Ao Engenheiro Florestal Joberto Veloso de Freitas, Professor da UFAM, Mestre pelo INPA e Doutor pela Universidade de Aberdeen, Escócia;

Ao Engenheiro Florestal Joésio Deoclécio Pierin Siqueira, Vice-Presidente da STCP, Professor Aposentado da UFPR, Mestre e Doutor pela UFPR;

Ao Engenheiro Florestal José Natalino Macedo Silva, Mestre pela UFPR e Doutor pela Universidade de Oxford;

Ao Engenheiro Agrônomo José Antônio Aleixo da Silva, Professor da UFRPE, Mestre pela UFV e Doutor pela University of Georgia;

Ao Engenheiro Florestal Paulo Luiz Contente de Barros, Professor Aposentado da UFRA, Mestre e Doutor pela UFPR;

Ao Engenheiro Florestal Sylvio Péllico Netto, Professor da UFPR, Mestre pela York-University e Doutor pela Albert-Ludwigs-Universität Freiburg.

Agradeço a todos os meus ex-alunos pela oportunidade de aprender ensinando e, em especial, aos estudantes do PET FLORESTAL/UFRA que, em 2017, na aula da saudade que homenageou os professores em processo de aposentadoria, presentearam-me com uma plaqueta — a qual foi incorporada em minha mente e que sempre me orientou na busca de um melhor aprimoramento —, apresentando a seguinte mensagem:

“Seu amor pela matemática nos mostrou a importância da Engenharia para a floresta. Agradecemos por nos mostrar que devemos ser autodidatas e que, por mais difícil que um assunto seja, temos que ter determinação e foco no aprendizado, pois em algum momento seremos capazes de aprender”.

Finalmente a todos aqueles que, direta ou indiretamente, contribuíram com a elaboração desta obra, sintam-se homenageados, pois foram muito importantes com suas sugestões e colaboração.



BIOGRAFIA DO AUTOR

Waldenei Travassos de Queiroz nasceu em Belém do Pará, no ano de 1950. É graduado em Engenharia Florestal e Mestre em Manejo Florestal pela Universidade Federal do Paraná (UFPR) e doutor em Estatística e Experimentação Agronômica pela Escola Superior de Agricultura “Luiz de Queiroz” (USP). Em 1974, iniciou suas atividades profissionais no Projeto RadamBrasil. Em 1975 e 1976, foi contraparte Nacional do Setor de Inventário Florestal do Projeto BRA-45 FAO/IBDF. É Professor Titular aposentado da Universidade Federal Rural da Amazônia (UFRA). Em 1979, foi membro da comissão do Instituto Brasileiro de Desenvolvimento Florestal (IBDF) para traçar diretrizes básicas para o estabelecimento de um Sistema Nacional de Avaliação de Recursos Naturais do Brasil. Em 1997, foi consultor do Programa de Desenvolvimento das Nações Unidas (PNUD) em Estatística e Planejamento de Inventários Florestais no Programa de Cooperação Técnica ao Zoneamento Ecológico Econômico do Estado de Rondônia (Projeto BRA/94/007). Foi Pró-reitor de Pesquisa e de Pós-Graduação e Vice-Reitor da UFRA. De 1988 a 2002, foi professor de Inventário Florestal e Análise Multivariada do Curso de Pós-Graduação em Ciências Florestais Tropicais do Instituto Nacional de Pesquisa da Amazônia (INPA). Entre 2009 e 2010, foi consultor da Organização das Nações Unidas para a Agricultura e Alimentação (FAO), junto ao Serviço Florestal Brasileiro (SFB) do Ministério do Meio Ambiente (MMA), para o planejamento do Sistema de Amostragem do Inventário Florestal Nacional do Brasil. Publicou os livros “Técnicas de Amostragem em Inventário Florestal nos Trópicos”, em 1998; “Amostragem em Inventário Florestal”, em 2012; “Análise Multivariada em Inventário Florestal Contínuo”, em 2021. É membro do Conselho Editorial e Científico da Revista Ciência Florestal da Universidade Federal de Santa Maria (UFSM) e também membro da Comissão Editorial da Revista Árvore da Universidade Federal de Viçosa (UFV).



PREFÁCIO

O autor da presente publicação apresenta à comunidade científica e, especialmente, aos discentes dos cursos de graduação e pós-graduação em ciências florestais, toda a sua experiência adquirida durante os 41 anos de efetivo exercício do magistério superior, dispensados ao ensino dos alunos de graduação da Universidade Federal Rural da Amazônia - UFRA, bem como aos discentes de pós-graduação dos cursos da UFRA e de outras instituições de ensino e de pesquisa da região Amazônica, tais como o Instituto Nacional de Pesquisas da Amazônia - INPA.

Durante sua jornada, o professor Waldenei Travassos de Queiroz se dedicou aos estudos da mensuração de recursos florestais e à quantificação das variações das mais diversas variáveis dendrométricas, coletadas e analisadas em estudos realizados no âmbito da ciência florestal, especialmente nas florestas tropicais de terra firme da Amazônia brasileira. Esses estudos proporcionaram ao autor um vasto conhecimento das técnicas de análises estatísticas empregadas em dados coletados nas pesquisas florestais.

O Professor Waldenei se dedicou, inicialmente, aos estudos das diferentes técnicas de análise de dados amostrais coletados nos inventários florestais, especialmente os realizados na Amazônia que culminaram, em 1998, com a elaboração e publicação de seu primeiro livro, intitulado “TÉCNICAS DE AMOSTRAGEM EM INVENTÁRIO FLORESTAL NOS TRÓPICOS”. Nessa obra o autor discorre de forma didática sobre o tema “AMOSTRAGEM”, apresentando e ilustrando com exemplos a aplicação dos processos de amostragem mais utilizados na época pela comunidade científica florestal.

Em 2012, verificando a necessidade de se aprofundar mais no tema “AMOSTRAGEM”, haja vista a demanda que se apresentava, especialmente, na realização pelo Serviço Florestal Brasileiro (SFB) do Inventário Florestal Nacional (IFN), publicou o seu segundo livro, “AMOSTRAGEM EM INVENTÁRIO FLORESTAL”, o qual abordou tópicos complementares ao seu primeiro livro, apresentando os capítulos de “Estimativas por Razão”, “Estimativas por Regressão” e as diversas alternativas de análises do processo de “Amostragem por Conglomerado” em diferentes estágios.

A exigência legal do monitoramento das florestas sob Planos de Manejo Florestal Sustentável-PM-FS, por meio de Inventários Florestais Contínuos (IFC), fez surgir a necessidade de aprofundar os estudos acerca do tema IFC nos livros textos. Esse tema é tratado apenas nos tópicos referentes ao “Método de Análise de Parcelas Subdivididas” e no “Modelo de Amostragem com Substituição Parcial”. Portanto, surge a necessidade de se aprofundar os estudos sobre os métodos de análise de experimentos com medidas repetidas no tempo, principalmente para contribuir com a análise dos dados do Inventário Florestal Nacional. Assim, em 2021, o autor elaborou e publicou o seu terceiro livro, “ANÁLISE MULTIVARIADA EM INVENTÁRIO FLORESTAL CONTÍNUO”.

Agora, o autor apresenta seu mais recente livro, “UMA INTRODUÇÃO À EXPERIMENTAÇÃO FLORESTAL”, com o objetivo de oferecer uma contribuição científica e suprir a lacuna bibliográfica na análise de dados oriundos de medições de experimentos florestais. O livro não é apenas um manual ou guia prático, pois apresenta os diferentes tópicos em 11 capítulos com profundidade para embasar os leitores em suas atividades profissionais e de pesquisa.

Embora o autor tenha intitulado o livro como “UMA INTRODUÇÃO” à experimentação florestal, ele apresenta um leque de conteúdos de grande abrangência que engloba uma gama enorme de assuntos, os quais podem ser utilizados nas mais diferentes situações na experimentação.

O livro engloba as noções básicas de estatística e experimentação, diversos testes de comparações múltiplas e de contrastes de médias, todos os delineamentos básicos e suas variações, envolvendo parcelas subdivididas e experimentos fatoriais, como, também, análise de grupos de experimentos, blocos incompletos equilibrados, experimentos reticulados ou látices quadrados e por fim a metodologia de superfície de resposta.

Finalmente, para tornar ainda mais completa a obra, o autor apresenta um Apêndice com todos os 39 “PROGRAMAS SAS”, que foram elaborados e utilizados nos exemplos de aplicação dos assuntos tratados.

É justo afirmar que a presente obra se constitui em uma importante contribuição para preencher uma lacuna no campo da experimentação e pesquisa florestal do país.

Prof. Dr. Engenheiro Florestal Paulo Luiz Contente de Barros
Professor Aposentado da Universidade Federal Rural da Amazônia

APRESENTAÇÃO

Este livro contém a minha experiência no ensino de graduação e de pós-graduação em Estatística Experimental na Universidade Federal Rural da Amazônia (UFRA), no exercício ativo da docência. Este livro nada mais será do que “um pingô d’água no oceano”, dentro do mar diversificado de alternativas que caracterizam o ramo da Estatística Experimental. O objetivo dele é contribuir para minimizar as dificuldades enfrentadas pelos estudantes e profissionais da engenharia florestal que necessitam utilizar métodos estatísticos na análise de experimentos.

O padrão científico das informações coletadas, a competitividade e a sustentabilidade, atualmente, são características fundamentais para as atividades florestais, absorvendo continuamente os avanços tecnológicos e inovações, o que está intimamente relacionado com os conhecimentos científicos dos pesquisadores envolvidos. A atuação profissional hoje requer inteligência aplicada no planejamento, execução e condução dos delineamentos experimentais, pois a exatidão das informações obtidas constitui o ponto essencial do problema.

Uma pesquisa científica, estatisticamente bem planejada, consiste em definir as etapas a serem cumpridas, as quais dependem de um perfeito entendimento entre o pesquisador e o biometrista florestal. A boa formação dos nossos estudantes de Engenharia Florestal em Estatística Experimental é uma condição obrigatória para o sucesso da ciência florestal.

O autor, na busca do aperfeiçoamento desta publicação e para melhor atender aos anseios das comunidades acadêmica e científica, coloca-se à disposição e agradecido pelo recebimento de críticas e sugestões.

Engenheiro Florestal Waldenei Travassos de Queiroz
Professor Aposentado da Universidade Federal Rural da Amazônia



SUMÁRIO

CAPÍTULO 1 - NOÇÕES BÁSICAS DE ESTATÍSTICA E EXPERIMENTAÇÃO...	20
1.1 - CONSIDERAÇÕES GERAIS.....	21
1.2 - VARIÁVEIS E DISTRIBUIÇÕES PROBABILÍSTICAS	25
1.2.1 - Esperança matemática ou valor médio.....	26
1.2.2 - Medidas de dispersão	26
1.2.3 - Amostragem com e sem reposição.....	27
1.2.4 - Distribuição Binomial.....	32
1.2.5 - Distribuição de Poisson	34
1.2.6 - Distribuição Normal.....	35
1.2.7 - Distribuição de Qui-Quadrado χ^2	37
1.2.8 - Distribuição “F” de Snedecor	39
1.2.9 - Distribuição “t” de Student	39
1.2.10 - Conceitos sobre inferência estatística e distribuição por amostragem	40
1.2.11 - Teste do valor médio versus um valor padrão populacional.....	41
1.2.12 - Comparação dos valores médios para duas populações independentes.....	43
1.2.13 - Comparação dos valores médios para duas populações dependentes. Teste “t” pareado.....	44
1.3 - EXIGÊNCIAS DO MODELO MATEMÁTICO NA APLICAÇÃO DA ANÁLISE DE VARIÂNCIA.....	45
1.3.1 - Teste de homogeneidade de variâncias pela razão máxima de Hartley.....	45

1.3.2 - Homogeneidade de variâncias pelo teste de Bartlett.....	46
1.3.3 - Homogeneidade de variâncias pelo teste de Cochran	47
1.3.4 - Teste de normalidade pelo Qui-Quadrado χ^2	47
1.3.5 - Teste de aderência de Kolmogorov-Smirnov (KS).....	50
1.3.6 - Teste de aderência de Lilliefors	53
1.3.7 - Teste de aderência de Cramér-Von Mises.....	55
1.3.8 - Teste de aderência de Anderson-Darling.....	56
1.3.9 - Teste de Shapiro-Wilk (1965).....	58
1.4 BIBLIOGRAFIA.....	61
 CAPÍTULO 2 - DELINEAMENTO INTEIRAMENTE AO ACASO	62
2.1 - INTRODUÇÃO.....	63
2.2 - PRESSUPOSIÇÕES BÁSICAS PARA A ANÁLISE DE VARIÂNCIA	63
2.3 - DESDOBRAMENTO DA SOMA DE QUADRADOS TOTAL E O TESTE F	65
2.4 - COMPONENTES DE VARIÂNCIA DO MODELO INTEIRAMENTE AO ACASO	72
2.5 - TRANSFORMAÇÕES DE DADOS NO CASO DE HETEROGENEIDADE DE VARIÂNCIAS	82
2.6 - ANÁLISE DE UM ENSAIO INTEIRAMENTE AO ACASO COM K AMOSTRAS POR PARCELA	87
2.7 - ANÁLISE DE COVARIÂNCIA	91
2.7.1 - Análise de covariância para o ensaio inteiramente ao acaso	93
2.8 - BIBLIOGRAFIA	102
 CAPÍTULO 3 - TESTES DE COMPARAÇÕES MULTIPLAS.....	104
3.1 - CONTRASTES DE MÉDIAS.....	105
3.2 - MÉTODO DE COMPARAÇÃO MÚLTIPLA DE TUKEY.....	112
3.3 - MÉTODO DE SCHEFFÉ.....	117
3.4 - MÉTODO DE DUNCAN.....	121
3.5 - MÉTODO SNK (STUDENT, NEWMAN, KEULS)	123
3.6 - MÉTODO DE DUNNETT	124
3.7 - MÉTODO DE BONFERRONI	125
3.8 - COMPARAÇÕES ENTRE OS TESTES DE TUKEY, SCHEFFÉ, BONFERRONI E DUNCAN.....	128
3.9 - CÁLCULO DO NÚMERO DE REPETIÇÕES IDEAL	129
3.10 - BIBLIOGRAFIA	131

CAPÍTULO 4 - DELINEAMENTO CASUALIZADO EM BLOCOS	132
4.1 - INTRODUÇÃO.....	133
4.2 - ANÁLISE DE VARIÂNCIA (ANOVA) E COMPONENTES DE VARIÂNCIAS	136
4.3 - EFETIVIDADE DO DELINEAMENTO CASUALIZADO EM BLOCOS.....	145
4.4 - EXPERIMENTOS CASUALIZADOS EM BLOCOS COM TRATAMENTOS COMUNS...	146
4.4.1 - Análise de variância de ensaios casualizados em blocos com tratamentos comuns.....	147
4.5 - COMPONENTES DE VARIÂNCIA PARA O ENSAIO CASUALIZADO EM BLOCOS COM K AMOSTRAS POR PARCELA.....	155
4.6 - ANÁLISE DE EXPERIMENTOS DE ADUBAÇÃO.....	156
4.6.1 - Análise considerando a regressão polinomial.....	156
4.6.2 - Análise considerando a equação de Mitscherlich	160
4.7 - BIBLIOGRAFIA	164
 CAPÍTULO 5 - DELINEAMENTO EM QUADRADO LATINO	166
5.1 - INTRODUÇÃO.....	167
5.2 - DISTRIBUIÇÃO DOS TRATAMENTOS NAS UNIDADES EXPERIMENTAIS	168
5.3 - ANÁLISE DE VARIÂNCIA	169
5.4 - BIBLIOGRAFIA	175
 CAPÍTULO 6 - EXPERIMENTOS FATORIAIS	176
6.1 - CONSIDERAÇÕES GERAIS.....	177
6.2 - ANÁLISE DE VARIÂNCIA DOS EXPERIMENTOS COM DOIS FATORES.....	181
6.3 - EXPERIMENTOS FATORIAIS HIERÁRQUICOS	187
6.3.1 - Análise de variância de experimentos hierárquicos em dois estágios	188
6.4 - CONFUNDIMENTO	195
6.4.1 - Confundimento do fatorial 2^3	195
6.4.2 - Confundimento do fatorial 3^3	199
6.4.3 - Confundimento do fatorial 3^2	202
6.5 - NOÇÕES SOBRE FATORIAIS FRACIONÁRIOS OU FRACIONADOS	203
6.5.1 - Construção da matriz de planejamento do experimento fatorial 2^k	204
6.5.2 - Planejamento de um fatorial 2^{k-p} e escolha das combinações a serem testadas.....	207
6.5.3 - Gerador e relação definidora	211
6.5.4 - Definição de resolução em um experimento fracionário	218
6.6 - EXPERIMENTOS FATORIAIS COM TRATAMENTOS ADICIONAIS.....	221
6.7 - BIBLIOGRAFIA	228

CAPÍTULO 7 - ANÁLISE DE GRUPOS DE EXPERIMENTO	230
7.1 - CONSIDERAÇÕES GERAIS	231
7.2 - ANÁLISE ESTATÍSTICA DE GRUPOS DE EXPERIMENTOS	231
7.2.1 - Análise de variância individual (por experimento ou local)	233
7.2.2 - Análise de variância conjunta para o grupo	234
7.3 - EXEMPLOS ILUSTRATIVOS	237
7.4 - BIBLIOGRAFIA	251
CAPÍTULO 8 - ENSAIOS COM PARCELAS SUBDIVIDIDAS (SPLIT- PLOT)	254
8.1 - EXPERIMENTOS EM PARCELAS SUBDIVIDIDAS	255
8.2 - EXPERIMENTOS EM FAIXAS “SPLIT BLOCK”	266
8.3 - EXPERIMENTOS COM PARCELAS SUBDIVIDIDAS NO TEMPO (SPLIT PLOT IN TIME)	276
8.4 - BIBLIOGRAFIA	282
CAPÍTULO 9 - DELINEAMENTOS EM BLOCOS INCOMPLETOS EQUILIBRADOS	284
9.1 - INTRODUÇÃO	285
9.2 - BLOCOS INCOMPLETOS BALANCEADOS (BIB)	287
9.3 - ANÁLISE ESTATÍSTICA DE BLOCOS INCOMPLETOS EQUILIBRADOS (BALANCEADOS)	293
9.3.1 - Análise intrablocos em blocos incompletos equilibrados (balanceados)	294
9.3.2 - Análise com recuperação da informação interblocos em blocos incompletos equilibrados (balanceados)	306
9.4 - BIBLIOGRAFIA	322
CAPÍTULO 10 - EXPERIMENTOS RETICULADOS OU LÁTICES QUADRADOS	324
10.1 - INTRODUÇÃO	325
10.2 - LÁTICES QUADRADOS (SQUARE LATTICES)	328
10.2.1 - Análise intrablocos de reticulados quadrados	331
10.2. 2 - Análise com recuperação da informação interblocos de látices quadrados	347
10.2. 3 - Repetição do delineamento em látice	365
10.3 - BIBLIOGRAFIA	366
CAPÍTULO 11 - METODOLOGIA DE SUPERFÍCIES DE RESPOSTA	368
11.1 - ANÁLISE DE SUPERFÍCIE DE RESPOSTA DE UM ENSAIO 3×3	369

11.1.1 - Estimativas dos parâmetros da regressão.....	369
11.1.2 - Análise de variância para o modelo 11.1.1.2.....	373
11.1.3 - Cálculo das doses economicamente aconselháveis.....	374
11.1. 4 - Uso da análise canônica na otimização de modelos na forma quadrática.....	376
11.1.5 - Exemplo de um fatorial 3×3.....	380
11.2 - ANÁLISE DE SUPERFÍCIE DE RESPOSTA DE UM ENSAIO FATORIAL 3×3×3	387
11.2.1 - Estimativas dos parâmetros da regressão.....	387
11.2.2 - Doses econômicas aconselháveis	389
11.2.3 - Um exemplo de aplicação de um fatorial 3×3×3	390
11.3 - DELINEAMENTO COMPOSTO CENTRAL.....	395
11.3.1 - Delineamento composto central para $2^k = 8$ pontos fatoriais.....	396
11.4 - PLANEJAMENTO BOX-BEHNKEN	411
11.5 - EXEMPLO DE APLICAÇÃO DE UM FATORIAL $(\frac{1}{5})5^3$	417
11.6 - BIBLIOGRAFIA	422
 APÊNDICE A: DISTRIBUIÇÕES ESTATÍSTICAS.....	 424
APÊNDICE B: PROGRAMAS SAS	442

CAPÍTULO 1

NOÇÕES BÁSICAS DE ESTATÍSTICA E EXPERIMENTAÇÃO

Definições e conceitos sobre as variáveis aleatórias, assim como sobre as seguintes distribuições: Binomial, Poisson, Normal, Qui-quadrado, “F” de Snedecor e “t” de Student. Apresenta noções de estatística básica que são fundamentais para o entendimento das técnicas de análise empregadas nos estudos científicos, em especial na “Experimentação Florestal”, isto é, aborda as considerações gerais sobre o planejamento, execução, análise e interpretação dos resultados dos experimentos.

É feita também uma abordagem geral sobre inferência estatística e distribuição por amostragem, intervalo de confiança, teste do valor médio de uma população versus um valor padrão populacional, comparação dos valores médios para duas populações independentes e comparação dos valores médios para duas populações dependentes (teste “t” pareado). Trata das exigências do modelo matemático na aplicação da análise de variância — como a verificação de homogeneidade de variâncias pelos testes de Hartley, Bartlett e Cochran —, assim como da análise de normalidade pelos testes de Qui-Quadrado, Kolmogorov-Smirnov, Lilliefors, Cramér-Von Mises, Anderson-Darling e o Shapiro-Wilk.

1.1 CONSIDERAÇÕES GERAIS

Estatística é a matemática aplicada à análise de dados de observação. A Estatística Experimental estuda o planejamento, execução, análise e interpretação de experimentos, e se caracteriza como um trabalho planejado, obedecendo a determinados princípios básicos, visando comparar efeitos de tratamentos.

O procedimento geral adotado em uma pesquisa científica é formular hipóteses e testá-las a partir de um conjunto de observações realizadas em uma população sobre os parâmetros de interesse. Assim sendo, é imprescindível que o planejamento do experimento estabeleça o esquema da análise estatística sob a qual as hipóteses serão verificadas. Portanto, considerando que as hipóteses foram definidas e usando métodos de análise, torna-se imperativo que o planejamento do experimento e a análise dos dados estejam intimamente interligados.

Peres e Saldiva (1982) estabelecem que uma pesquisa científica, estatisticamente bem planejada, consiste em definir as etapas a serem cumpridas, as quais dependem de um perfeito entendimento entre o pesquisador e o estatístico.

A seguir, um exemplo das etapas a serem consideradas no planejamento de um experimento:

1. Enunciação do problema com a respectiva formulação das hipóteses

Em uma pesquisa científica, primeiramente são formuladas as hipóteses em termos científicos, e em seguida, respeitando a devida correspondência, as hipóteses são expressas em termos estatísticos. Nesse sentido, quando um pesquisador tem como objetivo estudar o efeito de vários tipos de adubos, os quais diferem pela quantidade de um determinado micronutriente, na verificação da produção de biomassa de uma determinada espécie florestal, tais formulações são necessárias. Para alcançar esse objetivo, o experimento pode ser planejado visando atender as seguintes finalidades:

- (a1) Comparar as médias correspondentes à produção de biomassa, para cada um dos tipos de adubos, no final da rotação das espécies florestais (hipótese de igualdade de médias ou hipótese de nulidade);
- (a2) Estabelecer, no caso de dosagem de adubação, uma relação funcional entre o crescimento médio das plantas e a quantidade do micronutriente.

2. Definição dos fatores (variáveis independentes) e seus respectivos níveis

Em um experimento, cujo objetivo fosse estudar o efeito na produção de madeira, mas considerando três tipos de podas em duas espécies florestais, há dois fatores: a poda com três níveis e a espécie com dois níveis. Nesse caso, o experimento envolve seis tratamentos, que são as combinações referentes aos níveis dos dois fatores.

Os fatores podem ser fixos ou aleatórios. O fator fixo ocorre quando os fatores são fixados e as conclusões abrangem apenas os níveis do fator testado no experimento. Já no caso dos fatores aleatórios, as conclusões são estendidas para a população dos níveis.

3. Definição da unidade experimental ou unidade de observação

Na maioria dos casos, na pesquisa em ciências agrárias, a unidade experimental é facilmente determinada, podendo ser uma planta ou um animal. Entretanto, às vezes, a unidade experimental pode ser um conjunto de plantas ou uma área plantada. O importante é que a escolha da unidade experimental tenha, como critério básico, o objetivo de buscar a maior homogeneidade possível, visando minimizar o erro experimental e assim detectar mais facilmente os efeitos dos tratamentos.

4. Definição das variáveis que serão medidas nas unidades de observação

Aplicados os tratamentos às unidades experimentais, as medidas realizadas constituem os valores das variáveis dependentes. A forma de obtenção da medida da variável é importante na precisão das observações, assim como a sua distribuição de probabilidade é fundamental para a escolha do método de análise estatística.

5. Estabelecimento dos procedimentos para alocação dos tratamentos (combinação de níveis de fatores) às unidades experimentais ou vice-versa

Este item sintetiza o que se denomina planejamento estatístico, pois se trata de regras que associam as unidades experimentais aos tratamentos e, por consequência, determinam os diferentes delineamentos experimentais. Em resumo, no delineamento experimental, os tratamentos, as unidades experimentais e o erro experimental são elementos prioritários. Os princípios de aleatorização (casualização), repetição e controle do local são determinantes para caracterizar os diversos planos experimentais. Pimentel Gomes (1984) considera importante, na pesquisa em ciências agrárias, os seguintes delineamentos experimentais:

Inteiramente casualizado;

Casualizado em blocos completos;

Quadrados latinos;

Blocos incompletos equilibrados;

Blocos incompletos parcialmente equilibrados;

Reticulados (ou látices) quadrados, cúbicos e retangulares;

Quadrados reticulados (ou látices), os quais não devem ser confundidos com os reticulados quadrados;

Delineamentos com parcelas subdivididas;

Casualizado em Blocos com alguns tratamentos comuns.

6. Análise estatística dos dados

Esta etapa trata da verificação das hipóteses formuladas no estudo científico, as quais são expressas em termos da relação entre a variável dependente e os fatores envolvidos.

7. Relatório Final

O relatório final deverá conter as informações inerentes aos itens anteriores, assim como as conclusões relativas às medidas de precisão das estimativas. Nele, deve haver também a importância em relação a outras pesquisas já realizadas, além de recomendações para montagem de projetos experimentais futuros no setor da pesquisa em estudo.

Salsburg (2009) escreve que Ronald Aylmer Fisher (1890-1962), o arquiteto da Estatística Experimental, bacharelou-se, em 1912, em Matemática pela Universidade de Cambridge. Em 1919, chefiou um laboratório de Estatística na Estação Experimental de Agricultura de Rothamsted, Inglaterra, onde permaneceu até 1933. Ronald Fischer nesse período se dedicou e estudou, incansavelmente, os problemas do planejamento e análise de pesquisas na área agrária. O grande fascínio de Fisher pelos problemas da pesquisa em ciências agrárias inspirou o desenvolvimento dos métodos de análise de delineamentos experimentais, a análise de variância, que, em conjunto com a análise de covariância, também de sua autoria, constitui-se no instrumental fundamental para a interpretação das análises dos resultados dos experimentos planejados.

O desenho do experimento ou delineamento experimental é caracterizado pela maneira de distribuir e designar os tratamentos às unidades experimentais ou parcelas. Padovani (2014) destaca que esses métodos, procedentes do cotidiano agrícola, tornaram-se universais e aplicáveis em todas as áreas de conhecimento: Medicina, Psicologia, Engenharia, Odontologia, Biologia, Ecologia, entre outras. Cita ainda que, como a formalização dos procedimentos ocorreu em um ambiente agrário, a origem dos termos técnicos da experimentação apresenta conotação agrícola. O termo parcela foi criado para designar a unidade de área usada no experimento. Atualmente, a palavra parcela tem um significado mais geral, pois, dependendo do experimento, pode ser um animal, uma planta, uma quantidade de sementes, uma faixa de plantação, uma pessoa, uma peça anatômica, um corpo de prova, entre várias outras possibilidades que podem ser utilizadas como unidade experimental.

Na análise de experimentos, os fatores em estudo são denominados como características sobre as quais há interesse em verificar a inferência sobre uma resposta do experimento. Os níveis de um fator constituem os tratamentos a serem pesquisados. Um fator é indicado como quantitativo quando seus níveis são referentes a quantidades (doses de adubação, intensidade de corte no manejo florestal). Um fator é definido qualitativo quando seus níveis são relativos a atributos (procedências de espécies, por exemplo).

Definidos os fatores e seus respectivos níveis, que serão designados como os tratamentos do estudo, o tipo e o formato da unidade experimental (parcela) e a variável dependente (resposta), torna-se também necessário definir o esquema de alocação dos tratamentos às unidades experimentais, isto é, como deve ser estruturado o delineamento experimental.

Fischer (1935) estabeleceu os princípios sobre os delineamentos: repetição, casualização e controle local. O princípio da casualização é o procedimento estatístico considerado como de maior contribuição à Estatística Experimental. O sorteio das unidades experimentais assegura a fidedignidade das conclusões, pois a casualização fortalece a garantia de que as parcelas (unidades experimentais), com características diferentes, tenham igual probabilidade de serem designadas para todos os tratamentos.

O delineamento inteiramente ao acaso, também denominado completamente casualizado ou totalmente casualizado, é o único representante da categoria dos experimentos que não utiliza o controle local, ou seja, utiliza apenas os princípios da repetição e da casualização. Portanto, é um planejamento apropriado para experimentos de laboratório, estufas, casas de vegetação, pequenas áreas de campo consideradas homogêneas em relação à variável resposta, entre outros.

Por outro lado, vale destacar que o delineamento inteiramente ao acaso não demanda qualquer restrição sobre a distribuição dos tratamentos às parcelas, ou seja, pressuposição de homogeneidade entre as unidades experimentais. Portanto resulta que a única fonte de variação nos dados é a advinda dos tratamentos. Então como todas as outras características são controladas, consequentemente, as demais causas de variação que possam influenciar o experimento serão fontes de erro experimental.

As variáveis controláveis são aquelas que fazem parte do contexto do problema, assim como as denominadas de tratamentos (prescritas pelo pesquisador). O erro experimental contempla a variação entre as unidades experimentais e independe dos tratamentos. As variáveis não controláveis são classificadas em duas formas:

- a) As covariáveis que podem ser observadas e usadas nas análises;
- b) As que originam o erro experimental e não podem ser observadas.

Os erros de mensuração, a forma incorreta de aplicar os tratamentos, assim como danos físicos não registrados nas parcelas, produzem influências no erro experimental, gerando viés nas comparações entre tratamentos.

Padovani (2014) apresenta os seguintes passos recomendados para o planejamento estatístico e execução de uma pesquisa científica:

- a) Enunciado claro do problema e formulação das hipóteses que serão estudadas. As hipóteses científica e estatística devem manter uma correspondência perfeita e o enunciado precisa se apresentar de maneira clara e objetiva;
- b) Indicação dos fatores (variáveis independentes, isto é, variáveis controladas pelo pesquisador) do estudo (a escolha dos fatores e seus respectivos níveis que constituirão os tratamentos);
- c) Definição da unidade experimental (parcela) que deve ser estabelecida no sentido de minimizar o erro experimental. Por outro lado, é lógico supor a existência de um intervalo ideal de tamanhos de parcela, em que a eficiência é máxima e o custo é mínimo;
- d) Designação das variáveis (respostas), as quais serão medidas na unidade experimental. É importante observar que a distribuição probabilística associada à variável resposta é essencial para a escolha do método de análise estatística;
- e) Indicação das regras e procedimentos pelos quais os diferentes tratamentos (combinação de níveis de fatores) serão atribuídos às unidades experimentais (processo de casualização ou aleatorização);
- f) Análise estatística dos dados com o objetivo de verificar as hipóteses estabelecidas no planejamento da pesquisa;
- g) Descrição dos resultados analíticos com as medidas de precisão das estimativas e o respectivo nível de significância, assim como apresentação de suas interpretações inferenciais.

As duas situações mais comuns, no contexto da experimentação em ciências agrárias, que envolvem parcelas homogêneas e heterogêneas, são:

- i) Delineamento Inteiramente Casualizado (DIC), que consiste em alocar de forma inteiramente ao acaso os tratamentos às unidades experimentais. Nesse caso, a exigência é que as unidades experimentais sejam homogêneas;
- ii) Delineamento Casualizado em Blocos (DCB), que agrega, quando o conjunto é heterogêneo, grupos homogêneos (blocos) de unidades experimentais, em que os tratamentos são alocados casualmente às unidades experimentais dentro dos blocos.

A análise de variância (ANOVA) consiste na decomposição da variação total de um material heterogêneo em partes atribuídas a causas conhecidas e independentes, assim como em uma porção residual de origem desconhecida e de natureza aleatória. A ANOVA é desenvolvida em duas etapas. A primeira consiste em obter os componentes de variância e a segunda corresponde a proceder ao teste de significância, assim como apresenta a seguinte classificação:

- i) Análise de variância simples: há apenas um critério de classificação dos dados;
- ii) Análise de variância múltipla: há mais de um critério de classificação dos dados.

A aplicação mais simples da análise de variância, para I grupos ou tratamentos, consiste em testar a hipótese de igualdade de médias ou hipótese de nulidade, a qual é escrita quando o efeito de tratamento é fixo, da seguinte forma: $H_0 : m_1 = m_2 = \dots = m_I$.

Por outro lado, se o efeito de tratamentos é aleatório, então H_0 apresenta o formato $H_0 : \sigma_t^2 = 0$, que verifica se a variância do efeito de tratamento é nula. Na realização de uma análise de variância, em que os tratamentos são amostras casuais tiradas da mesma população normal de média m e variância σ^2 , as estimativas de σ^2 são independentes.

1.2. VARIÁVEIS E DISTRIBUIÇÕES PROBABILÍSTICAS

Diz-se que X é uma variável aleatória discreta se os valores possíveis de X (contradomínio) forem finitos ou infinito enumerável. Por outro lado, se o contradomínio de X for um intervalo ou um conjunto de intervalos, então X é uma variável aleatória contínua.

Sendo $X_1, X_2, \dots, X_n$ os valores observados de uma amostra aleatória de uma variável X , então $y = H(X_1, X_2, \dots, X_n)$ é uma estatística. São exemplos de estatísticas os estimadores da média e da variância:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad \hat{\sigma}^2 = \frac{1}{n-1} \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right]$$

1.2.1. Esperança matemática ou valor médio

Havendo uma variável aleatória discreta X , que assume os valores $X_1, X_2, \dots, X_n$, com as respectivas probabilidades $p_1, p_2, \dots, p_n$, define-se como esperança matemática de X , representada por $E[X]$:

$$E[X] = \sum_{i=1}^n X_i p_i, \text{ em que } \sum_{i=1}^n p_i = 1.$$

No caso de uma variável aleatória contínua X , então a esperança matemática é dada por:

$$E[X] = \int_{-\infty}^{\infty} X f(X) dx \quad \int_{-\infty}^{\infty} f(X) dX = 1$$

$f(X)$: é a função densidade da variável aleatória X .

Propriedades da esperança matemática:

$$E[c] = c \text{ (constante);}$$

$$E[cX] = cE[X];$$

$$E[X \pm Y] = E[X] \pm E[Y];$$

$$E[X \pm c] = E[X] \pm c;$$

$$E[XY] = E[X] \times E[Y]: \text{ se as variáveis } X \text{ e } Y \text{ forem independentes.}$$

1.2.2. Medidas de dispersão

A variância de uma variável aleatória X é dada por:

$$V(X) = E[(X - E[X])^2] = E[(X - \mu)^2]$$

$$\text{Para variável discreta: } V(X) = \sum_{i=1}^N (X_i - \mu)^2 p(x)$$

$$\text{Para variável contínua: } V(X) = \int_{-\infty}^{+\infty} (X - \mu)^2 f(x) dx$$

$$\text{a) } V(aX) = a^2 V(X);$$

$$\text{b) } V(X + Y) = V(X) + V(Y) + 2Cov(X, Y);$$

$$\text{c) } V(X - Y) = V(X) + V(Y) - 2Cov(X, Y).$$

$Cov(X, Y)$: covariância entre as duas variáveis aleatórias X e Y .

Variância populacional:

$$V(X) = \sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$$

Variância amostral:

$$\hat{V}(X) = \hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

Desvio padrão amostral:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2}$$

Fórmula mais prática para calcular a variância:

$$V(X) = E[(X - E[X])^2] = E[(X - \mu)^2] = E[X^2 - 2X\mu + \mu^2]$$

$$V(X) = E[X^2] - E[2X\mu] + E[\mu^2]$$

$$V(X) = E[X^2] - 2\mu E[X] + \mu^2$$

$$V(X) = E[X^2] - 2\mu \times \mu + \mu^2$$

$$V(X) = E[X^2] - \mu^2 \rightarrow E[X^2] = V(X) + \mu^2$$

A covariância entre duas variáveis aleatórias X e Y é definida como:

$$Cov(X, Y) = E\{(X - E[X])(Y - E[Y])\} = E[XY] - E[X]E[Y]$$

$$Cov(X, Y) = E[XY] - \mu_X \mu_Y$$

A variância de uma função linear $Y = a_1 X_1 + a_2 X_2 + \dots + a_n X_n$ de variáveis aleatórias é dada, por:

$$V(Y) = a_1^2 V(X_1) + a_2^2 V(X_2) + \dots + a_n^2 V(X_n) + 2a_1 a_2 Cov(X_1, X_2) + \dots + 2a_{n-1} a_n Cov(X_{n-1}, X_n)$$

Considerando apenas as variáveis X_1 e X_2 :

$$V(Y) = a_1^2 V(X_1) + a_2^2 V(X_2) + 2a_1 a_2 Cov(X_1, X_2)$$

Se X_1 e X_2 são independentes, então $Cov(X_1, X_2) = 0$, tem-se:

$$V(Y) = a_1^2 V(X_1) + a_2^2 V(X_2)$$

1.2.3. Amostragem com e sem reposição

Seja N o número de observações de uma população e n o tamanho da amostra, então há as seguintes definições:

a) No caso de a seleção ocorrer com reposição haverá N^n amostras de tamanho n ;

b) No caso de a seleção ocorrer sem reposição, haverá a combinação $\binom{N}{n}$ de amostras de tamanho n .

Sendo $X_1, X_2, \dots, X_n$ uma amostra aleatória e independente de n observações de uma população com uma média μ e variância σ^2 , em que $\bar{X}$ é o estimador da média populacional:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Como a amostra é aleatória e cada observação tem a mesma distribuição da população:

$$E[X_i] = \mu \quad V(X_i) = \sigma^2$$

$$E[\bar{X}] = E\left[\frac{\sum_{i=1}^n X_i}{n}\right] = \frac{1}{n} \sum_{i=1}^n E[X_i] = \frac{1}{n} \sum_{i=1}^n \mu = \frac{n\mu}{n} = \mu$$

$$V(\bar{X}) = V\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \sum_{i=1}^n V\left(\frac{X_i}{n}\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Quando $X \sim N(\mu, \sigma^2)$, então $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

No caso de populações finitas:

Se $X \sim N(\mu, \sigma^2)$, então $\bar{X} \sim N\left[\mu, \frac{\sigma^2}{n} \left(\frac{N-n}{N-1}\right)\right]$.

Exercício 1.2.3.1. Considerando uma população com $N = 5$ observações $[2, 4, 6, 8, 10]$ e o procedimento de amostragem com reposição, tem-se as $N^n = 5^2 = 25$ como possíveis amostras aleatórias de tamanho $n = 2$. Assim, obter o valor médio, a variância das observações e a variância da média.

$$\left\{ \begin{array}{ccccc} (2,2) & (2,4) & (2,6) & (2,8) & (2,10) \\ (4,2) & (4,4) & (4,6) & (4,8) & (4,10) \\ (6,2) & (6,4) & (6,6) & (6,8) & (6,10) \\ (8,2) & (8,4) & (8,6) & (8,8) & (8,10) \\ (10,2) & (10,4) & (10,6) & (10,8) & (10,10) \end{array} \right\}$$

a) Valores das estimativas das médias ($\bar{X}$) por amostra de tamanho $n = 2$:

$$\left\{ \begin{array}{ccccc} (2) & (3) & (4) & (5) & (6) \\ (3) & (4) & (5) & (6) & (7) \\ (4) & (5) & (6) & (7) & (8) \\ (5) & (6) & (7) & (8) & (9) \\ (6) & (7) & (8) & (9) & (10) \end{array} \right\}$$

$$\sum_{i=1}^{25} \bar{X}_i = 150,0 \quad \sum_{i=1}^{25} \bar{X}_i^2 = 1.000,0$$

b) Distribuição de probabilidade da variável aleatória ($\bar{X}$):

$\bar{X}$	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0	10,0
$P(\bar{X})$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{5}{25}$	$\frac{4}{25}$	$\frac{3}{25}$	$\frac{2}{25}$	$\frac{1}{25}$

c) Calcular a média populacional (μ), a variância populacional (σ^2) e a variância da média [$V(\bar{X}) = \frac{\sigma^2}{n}$]

$$\mu = 2 \times \frac{1}{25} + 3 \times \frac{2}{25} + 4 \times \frac{3}{25} + 5 \times \frac{4}{25} + 6 \times \frac{5}{25} + 7 \times \frac{4}{25} + 8 \times \frac{3}{25} + 9 \times \frac{2}{25} + 10 \times \frac{1}{25} = 6$$

Ou:

$$\mu = \frac{\sum_{i=1}^{25} \bar{X}_i}{25} = \frac{150}{25} = 6$$

Por outro lado:

$$\mu = \frac{1}{5} \sum_{i=1}^5 X_i = \frac{2+4+6+8+10}{5} = \frac{30}{5} = 6$$

d) Variância populacional:

$$V(X) = E[X^2] - \mu^2 = (2^2 \times \frac{1}{5} + 4^2 \times \frac{2}{5} + 6^2 \times \frac{3}{5} + 8^2 \times \frac{4}{5} + 10^2 \times \frac{1}{5}) - 6^2$$

$$V(X) = \frac{1}{5} (4+16+36+64+100) - 36^2 = \frac{220}{5} - 36 = 8$$

Ou:

$$V(X) = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} \rightarrow V(X) = \frac{\sum_{i=1}^5 X_i^2 - \frac{(\sum_{i=1}^5 X_i)^2}{N}}{N} = \frac{220 - \frac{30^2}{5}}{5} = \frac{40}{5} = 8$$

e) Variância da média

$$V(\bar{X}) = E(\bar{X}^2) - \mu^2 = (2^2 \times \frac{1}{25} + 3^2 \times \frac{2}{25} + 4^2 \times \frac{3}{25} + 5^2 \times \frac{4}{25} + 6^2 \times \frac{5}{25} + 7^2 \times \frac{4}{25} + 8^2 \times \frac{3}{25} + 9^2 \times \frac{2}{25} + 10^2 \times \frac{1}{25}) - 6^2$$

$$V(\bar{X}) = \frac{1}{25}(4+18+48+100+180+196+192+162+100)-36 = \frac{348}{25} - 36 = \frac{1.000,0}{25} - 36 = 4$$

Alternativamente:

$$V(\bar{X}) = \frac{\sum_{i=1}^{25} \bar{X}_i^2 - \frac{(\sum_{i=1}^{25} \bar{X}_i)^2}{25}}{25} = \frac{1.000,0 - \frac{150^2}{25}}{25} = \frac{100}{25} = 4$$

Por outro lado:

$$V(\bar{X}) = \frac{V(X)}{n} = \frac{\sigma^2}{n} = \frac{8}{2} = 4$$

Exercício 1.2.3.2. Mostrar que:

$$E\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{n}\right] = \sigma^2$$

$$E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}\right] = \left(\frac{n-1}{n}\right)\sigma^2$$

$$E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right] = \sigma^2$$

Tem-se que: $E[X] = \mu$ e $E[X - \mu] = 0$.

Dado que:

$$V(X) = E[X^2] - \mu^2 \rightarrow E[X^2] = \sigma^2 + \mu^2$$

$$V(\bar{X}) = E[\bar{X}^2] - \mu^2 \rightarrow E[\bar{X}^2] = \frac{\sigma^2}{n} + \mu^2$$

a)
$$E\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{n}\right] = \sigma^2$$

$$E\left[\sum_{i=1}^n (X_i - \mu)^2\right] = \sum_{i=1}^n E[(X_i - \mu)^2] = \sum_{i=1}^n E[X_i^2 - 2X_i\mu + \mu^2] = \sum_{i=1}^n E[X_i^2] - 2\mu \sum_{i=1}^n E[X_i] + \sum_{i=1}^n E[\mu^2]$$

Dado que:

$$\sum_{i=1}^n E[\mu^2] = n\mu^2 \quad \sum_{i=1}^n E[X_i] = n\mu \quad \sum_{i=1}^n E[X^2] = n(\sigma^2 + \mu^2)$$

Logo:

$$E\left[\sum_{i=1}^n (X_i - \mu)^2\right] = n(\sigma^2 + \mu^2) - 2n\mu^2 + n\mu^2 = n\sigma^2$$

Tem-se, então, a fórmula da variância quando é conhecida a média populacional (μ):

$$E\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{n}\right] = \sigma^2$$

$$\text{b) } E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}\right] = \left(\frac{n-1}{n}\right)\sigma^2$$

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = E\left[\sum_{i=1}^n X_i^2 - 2\bar{X}\sum_{i=1}^n X_i + \sum_{i=1}^n \bar{X}^2\right]$$

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = E\left[\sum_{i=1}^n X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right] \quad \sum_{i=1}^n X_i = n\bar{X}$$

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = E\left[\sum_{i=1}^n X_i^2 - n\bar{X}^2\right]$$

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = E\left[\sum_{i=1}^n X_i^2\right] - nE[\bar{X}^2]$$

Dado que:

$$V(X) = E[X^2] - \mu^2 \rightarrow E[X^2] = \sigma^2 + \mu^2$$

$$V(\bar{X}) = E[\bar{X}^2] - \mu^2 \rightarrow E[\bar{X}^2] = \frac{\sigma^2}{n} + \mu^2$$

Então:

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)$$

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2$$

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = (n-1)\sigma^2$$

Logo:

$$E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}\right] = \left(\frac{n-1}{n}\right)\sigma^2$$

$$c) E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right] = \sigma^2$$

Como $E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}\right] = \left(\frac{n-1}{n}\right)\sigma^2 \rightarrow E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = (n-1)\sigma^2$, então:

$$E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right] = \sigma^2$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} : \text{é a variância amostral.}$$

1.2.4. Distribuição Binomial

É uma distribuição discreta com função de probabilidade dada por: $f(x) = C_n^x p^x q^{n-x}$, em que:

p : probabilidade de realização do acontecimento favorável;

q : probabilidade de realização do acontecimento contrário;

x : número de vezes que se realiza o acontecimento favorável;

n : número de tentativas.

Exercício 1.2.4.1. Em uma amostra de cinco árvores ($n = 5$), verificou-se a ocorrência de duas árvores ($x = 2$) com qualidade de fuste tipo1 (fuste reto, cilíndrico, bem configurado e sem deterioração aparente). A aplicação de um teste de aderência comprovou que a variável número de árvores com qualidade de fuste tipo1 segue a distribuição binomial. Obter as probabilidades de ocorrência para $x = 0, 1, 2, 3, 4, 5$.

Dado que $p = \frac{x}{n} = \frac{2}{5}$, então $q = 1 - p = \frac{3}{5}$.

a) Se $f(x) = C_n^x p^x q^{n-x}$, então a probabilidade de não ocorrerem árvores com qualidade fuste tipo1 é obtida por:

$$f(0) = C_5^0 \left(\frac{2}{5}\right)^0 \left(\frac{3}{5}\right)^5 = \frac{243}{3.125,0} = 0,07776$$

Conclui-se que a probabilidade de não ocorrerem árvores com qualidade de fuste tipo1 é igual a 7,78%.

b) A probabilidade de ocorrer apenas uma árvore com qualidade de fuste do tipo1 é obtida por:

$$f(1) = C_5^1 \left(\frac{2}{5}\right)^1 \left(\frac{3}{5}\right)^4 = \frac{810}{3.125,0} = 0,2592$$

Conclui-se que a probabilidade de ocorrer uma árvore apenas com qualidade de fuste tipo1 é igual a 25,92%.

c) A probabilidade de ocorrerem duas árvores com qualidade de fuste do tipo1 é dada por:

$$f(2) = C_5^2 \left(\frac{2}{5}\right)^2 \left(\frac{3}{5}\right)^3 = \frac{1.080}{3.125,0} = 0,3456$$

Conclui-se que a probabilidade de ocorrerem duas árvores com qualidade de fuste tipo1 é igual a 34,56%.

d) A probabilidade de ocorrerem três árvores com qualidade de fuste do tipo1 é dada por:

$$f(3) = C_5^3 \left(\frac{2}{5}\right)^3 \left(\frac{3}{5}\right)^2 = \frac{720}{3.125,0} = 0,2304$$

Conclui-se que a probabilidade de ocorrerem três árvores com qualidade de fuste tipo1 é igual a 23,04%.

e) A probabilidade de ocorrerem quatro árvores com qualidade de fuste tipo1 é dada por:

$$f(4) = C_5^4 \left(\frac{2}{5}\right)^4 \left(\frac{3}{5}\right)^1 = \frac{240}{3.125,0} = 0,0768$$

Conclui-se que a probabilidade de ocorrerem quatro árvores com qualidade de fuste tipo1 é igual a 7,68%.

f) A probabilidade de ocorrerem cinco árvores com qualidade de fuste tipo1 é dada por:

$$f(5) = C_5^5 \left(\frac{2}{5}\right)^5 \left(\frac{3}{5}\right)^0 = \frac{32}{3.125,0} = 0,01024$$

Conclui-se que a probabilidade de ocorrerem cinco árvores com qualidade de fuste tipo 1 é igual a 1,02%.

Observar que: $(p + q)^5 = \sum_{i=0}^n C_i^n (p)^i (q)^{n-i}$.

Verificação:

$$f(0) + f(1) + f(2) + f(3) + f(4) + f(5) = 1$$

$$0,0778 + 0,2592 + 0,3456 + 0,2304 + 0,0768 + 0,0102 = 1$$

1.2.5. Distribuição de Poisson

É uma distribuição discreta, cuja função de probabilidade é definida como:

$$f(x) = \frac{e^{-m} m^x}{x!}, \text{ em que } x \text{ pode assumir valores } x = 0, 1, \dots, \infty.$$

x : número de vezes em que se realiza o acontecimento;

m : média de ocorrência;

e : base do logaritmo natural.

Exercício 1.2.5.1. Em parcelas de tamanho igual a 0,25 ha, em um inventário de uma área florestal, ocorreram em média 20 árvores, considerando todas as espécies. Dado que a probabilidade de ocorrer uma árvore da família Meliaceae é de 10 %, e que, por meio de um teste de aderência, foi verificado que a variável resposta número de árvores ocorrentes desta família segue a distribuição de Poisson, obter as probabilidades de ocorrência na floresta de: nenhuma árvore, apenas uma árvore, três árvores, e cinco árvores da família Meliaceae.

Dado que: $f(x) = \frac{e^{-m} m^x}{x!}$ e $m = pn = 0,10 \times 20 = 2$, tem-se:

a) A probabilidade de não ocorrer árvore da família Meliaceae é dada por:

$$f(0) = \frac{e^{-2} \times 2^0}{0!} = \frac{1}{e^2} = 0,1353$$

Conclui-se que a probabilidade de não ocorrer árvore da família Meliaceae é de 13,53%.

b) A probabilidade de ocorrer apenas uma árvore da família Meliaceae é dada por:

$$f(1) = \frac{e^{-2} \times 2^1}{1!} = \frac{2}{e^2} = 0,2707$$

Conclui-se que a probabilidade de ocorrer apenas uma árvore da família Meliaceae é de 27,07%.

c) A probabilidade de ocorrerem três árvores da família Meliaceae é dada por:

$$f(3) = \frac{e^{-2} \times 2^3}{3!} = \frac{8}{6e^2} = 0,1804$$

Então, a probabilidade de ocorrerem três árvores da família Meliaceae é de 18,04%.

d) A probabilidade de ocorrerem cinco árvores da família Meliaceae é dada por:

$$f(5) = \frac{e^{-2} \times 2^5}{5!} = \frac{32}{120e^2} = 0,0361$$

Conclui-se que a probabilidade de ocorrerem cinco árvores da família Meliaceae é de 3,61%.

1.2.6. Distribuição normal

A distribuição normal também pode ser chamada de gaussiana e sua forma se assemelha à de um sino.

Diz-se que X é uma variável aleatória contínua com distribuição normal, se:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} \quad f(X) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} dX = 1$$

Tal que: $-\infty < X < \infty$, tal que μ e σ^2 são os parâmetros: média e variância.

Se $Z = \frac{X-\mu}{\sigma}$, então $f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2}$, em que $-\infty < Z < \infty$, tal que: $E(Z) = 0$ e $V(Z) = 1$.

Destacam-se as seguintes propriedades da distribuição normal:

- a) $f(X)$ é simétrica em relação à origem $X = \mu$ ou $f(Z)$ é simétrica em relação à origem $Z = 0$;
- b) $f(X)$ possui um ponto de máximo em $X = \mu$ ou $f(Z)$ apresenta um ponto de máximo para $Z = 0$. No caso de $f(X)$, a ordenada no ponto de máximo é igual a $\frac{1}{\sigma\sqrt{2\pi}}$, enquanto para $f(Z)$ é igual a $\frac{1}{\sqrt{2\pi}}$;
- c) $f(X)$ e $f(Z)$ tendem para zero quando X e Z tendem para $\pm\infty$;
- d) $f(X)$ possui dois pontos de inflexão, cujas abscissas são $\mu - \sigma$ e $\mu + \sigma$, assim como $f(Z)$ possui dois pontos de inflexão, cujas abscissas são -1 e $+1$.

Por outro lado, considerando o gráfico da curva normal, verifica-se que as áreas incluídas entre ($Z = -1$ e $Z = +1$) ou ($\mu - \sigma$ e $\mu + \sigma$), ($Z = -2$ e $Z = +2$) ou ($\mu - 2\sigma$ e $\mu + 2\sigma$) e ($Z = -3$ e $Z = +3$) ou ($\mu - 3\sigma$ e $\mu + 3\sigma$) são iguais, respectivamente, a 68,27 %, 95,45 % e 99,73 % da área total que é unitária. Essas informações são importantes para interpretar, principalmente em inventários florestais, a variabilidade das observações em uma amostragem.

O teorema da combinação linear mostra que a combinação linear de variáveis normais independentes é também uma variável normal.

Se X e Y são variáveis normais independentes, então $Z = a_1X + a_2Y$ é uma variável normal com média $\mu_z = a_1\mu_x + a_2\mu_y$ e variância $\sigma_z^2 = a_1^2\sigma_x^2 + a_2^2\sigma_y^2$.

Havendo $(X_1, X_2, \dots, X_n)$ uma amostra aleatória com média amostral $\bar{X}$ de uma população com média μ e variância σ^2 , no caso da distribuição por amostragem de $\bar{X}$, tem-se:

$$E(\bar{X}) = E\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{n\mu}{n} = \mu$$

$$V(\bar{X}) = V\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Exercício 1.2.6.1. Em um povoamento cultivado com 1000 árvores de *Bertholletia excelsa* (castanha do Brasil), foi observada a produção do número de castanhas (frutos) por ano de 30 árvores.

497	488	504	492	501	500	496	505	496	499
498	490	483	505	514	501	481	505	511	496
506	488	500	508	480	507	494	509	491	507

Os testes de aderência de Kolmogorov-Smirnov, Cramér-Von Mises, Anderson-Darling e Shapiro-Wilk aplicados, descritos nos itens 1.3.5, 1.3.7, 1.3.8 e 1.3.9, mostraram (tabela 1.2.6.1) que os dados da produção do número de castanhas (frutos) por árvore, ao nível de significância de $\alpha = 0,05$, são normalmente distribuídos.

Tabela 1.2.6.1. Testes de normalidade para o número de frutos.

Testes de Normalidade				
Testes	Estatísticas		p-valor	
Shapiro-Wilk	W	0,967915	Pr < W	0,4838
Kolmogorov-Smirnov	D	0,104351	Pr > D	> 0,1500
Cramér-Von Mises	W-Sq	0,045349	Pr > W-Sq	> 0,2500
Anderson-Darling	A-Sq	0,313084	Pr > A-Sq	> 0,2500

a) obter o número de árvores que produzem acima ou igual a 500 frutos por ano.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{14.952,0}{30} = 498,4$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}{n-1} = \frac{7.454.370,0 - \frac{14.952,0^2}{30}}{29} = 79,0759$$

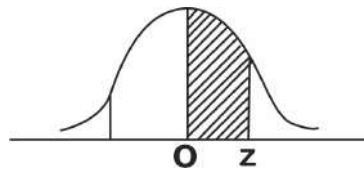
$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = 8,8925$$

Então:

$$P(X \geq 500) = \int_{500}^{\infty} \frac{1}{8,8925\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-498,4}{8,8925}\right)^2} dX$$

Dado que: $Z = \frac{X - \bar{X}}{\hat{\sigma}} = \frac{500 - 498,4}{8,8925} = 0,1799$

Da tabela A1, para $Z = 0,1799$, tem-se o valor 0,0714. Como:



Então: $f(Z) = \int_{0,1799}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} = 0,5 - 0,0714 = 0,4286$. É esperado que 429 árvores de castanheira tenham uma produção igual ou superior a 500 frutos por ano.

b) obter o número árvores que produzem 490 frutos ou menos por ano.

Então: $P(X \leq 490) = \int_{-\infty}^{490} \frac{1}{8,8925\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-498,4}{8,8925}\right)^2} dX$

Dado que $Z = \frac{X - \bar{X}}{\hat{\sigma}} = \frac{490 - 498,4}{8,8925} = -0,9446$

Da tabela A1, para $Z = -0,9446$, tem-se o valor 0,3289.

Então: $f(Z) = \int_{-\infty}^{-0,9446} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} = 0,5 - 0,3289 = 0,1711$. Logo é esperado que 171 árvores de castanheira tenham 490 frutos ou menos por ano.

1.2.7. Distribuição de Qui-Quadrado χ^2

A variável (χ^2), com k graus de liberdade, é definida como sendo a soma de k quadrados de normais padronizadas e independentes:

$$\chi_k^2 = \sum_{i=1}^k Z_i^2 = \sum_{i=1}^k \left(\frac{X_i - \mu}{\sigma}\right)^2$$

Tal que:

$$E(\chi_k^2) = k \quad \text{Var}(\chi_k^2) = 2k$$

Então $f(\chi_k^2)$ apresenta a seguinte distribuição de densidade:

$$f(\chi_k^2) = f(x) = \frac{x^{\frac{k}{2}-1} \times e^{-\frac{1}{2}x}}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} I_{(0,\infty)}(x) \quad (x = \chi_k^2)$$

k : graus de liberdade

$\Gamma(\frac{k}{2})$: função gama

A função gama é definida analiticamente como:

$$\Gamma(\alpha + 1) = \int_0^{\infty} x^{\alpha} e^{-x} dx$$

Em um experimento, a discrepância entre as frequências observadas e esperadas pode ser avaliada pela estatística qui-quadrado (χ^2):

$$\chi_v^2 = \frac{(o_1 - e_1)^2}{e_1} + \frac{(o_2 - e_2)^2}{e_2} + \dots + \frac{(o_k - e_k)^2}{e_k} = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$$

Exercício 1.2.7.1. De acordo com Spiegel (1972), nas experiências de Mendel com 556 ervilhas, foram observados os seguintes resultados: 315 ervilhas redondas e amarelas, 108 redondas e verdes, 101 enrugadas e amarelas e 32 enrugadas e verdes. De acordo com a sua teoria da hereditariedade, os números deveriam estar na proporção 9: 3: 3: 1. Isto é:

$$\frac{9}{16} : \text{ervilhas amarelas lisas} \rightarrow \frac{9}{16} \times 556 = 312,75$$

$$\frac{3}{16} : \text{ervilhas amarelas rugosas} \rightarrow \frac{3}{16} \times 556 = 104,25$$

$$\frac{3}{16} : \text{ervilhas verdes lisas} \rightarrow \frac{3}{16} \times 556 = 104,25$$

$$\frac{1}{16} : \text{ervilhas verdes rugosas} \rightarrow \frac{1}{16} \times 556 = 34,75$$

$$\chi_v^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} = \frac{(315-312,75)^2}{312,75} + \frac{(108-104,25)^2}{104,25} + \frac{(101-104,25)^2}{104,25} + \frac{(32-34,75)^2}{34,75} = 0,470$$

Dado que $v = 4 - 1 = 3$, então, de acordo com a tabela A5, tem-se: $\chi_{3; 0,05}^2 = 7,81$. Logo se conclui que a teoria e a experiência conduzida estão em concordância.

1.2.8. Distribuição “F” de Snedecor

A distribuição F é a razão entre duas variáveis independentes distribuídas como χ^2 , podendo possuir ou não o mesmo número de graus de liberdade:

$$F_{(m,p)} = \frac{\frac{\chi_m^2}{m}}{\frac{\chi_p^2}{p}} = \frac{\chi_m^2}{\chi_p^2} \times \frac{p}{m}$$

A estatística “ F ” tem distribuição de Snedecor com m graus de liberdade no numerador e p graus de liberdade no denominador e apresentando a seguinte função de densidade:

$$f(x) = \frac{\Gamma(\frac{m+p}{2})}{\Gamma(\frac{m}{2}) \times \Gamma(\frac{p}{2})} \left(\frac{m}{p}\right)^{\frac{m}{2}} \times \frac{x^{\frac{m-2}{2}}}{[1 + (\frac{m}{p})x]^{\frac{m+p}{2}}} I_{(0,\infty)}(x) \quad (x = F)$$

1.2.9. Distribuição “t” de Student

Dada uma amostra selecionada de uma população normal com média μ e variância σ^2 , sendo

$s^2 = \frac{1}{n-1} [\sum_{i=1}^n (X_i - \bar{X})^2]$ a variância amostral e $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ a média amostral, como s^2 tem distribuição $\frac{\sigma^2}{n-1} \times \chi_{n-1}^2$ com média zero e variância $\frac{2\sigma^4}{n-1}$, tem-se:

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \times \frac{\sigma}{s} = \frac{N(0,1)}{\frac{s}{\sigma}} = \frac{N(0,1)}{\sqrt{\frac{\chi_{n-1}^2}{n-1}}}$$

que é conhecida como distribuição “t” de Student com $n-1$ graus de liberdade, com média zero e variância igual a $\frac{k}{k-2}$, sendo k o número de graus de liberdade e apresentando a seguinte função de densidade:

$$f(t) = \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \times \frac{1}{\sqrt{k\pi}} \times \frac{1}{(1 + \frac{t^2}{k})^{\frac{k+1}{2}}}$$

Intervalo de confiança:

$$P(-t_{\frac{\alpha}{2}} \leq t \leq t_{\frac{\alpha}{2}}) = 1 - \alpha \quad \rightarrow \quad (-t_{\frac{\alpha}{2}} \leq \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \leq t_{\frac{\alpha}{2}}) = 1 - \alpha$$

Resultando:

$$P(\bar{x} - t_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}) = 1 - \alpha$$

Intervalo de confiança para a variância de uma população normal:

Havendo a variável aleatória X com distribuição normal, média μ e variância σ^2 , então:

$$\chi_{n-1}^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$P(\chi_{\inf}^2 \leq \chi^2 \leq \chi_{\sup}^2) = 1 - \alpha$$

$$P[\chi_{\inf}^2 \leq \frac{(n-1)s^2}{\sigma^2} \leq \chi_{\sup}^2] = 1 - \alpha$$

$$P(\frac{(n-1)s^2}{\chi_{\inf}^2} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi_{\sup}^2}) = 1 - \alpha$$

1.2.10. Conceitos sobre inferência estatística e distribuição por amostragem

a) A estimação consiste em produzir inferências sobre a população a partir dos resultados extraídos de amostra selecionada aleatoriamente.

a1) Estimação por ponto: a partir da amostra se infere sobre o valor de um determinado parâmetro populacional, como no exemplo a seguir:

$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$: é uma estatística que produz uma estimativa por ponto do parâmetro populacional $\mu = \frac{\sum_{i=1}^N X_i}{N}$.

a2) Estimação por intervalo: calcula-se um intervalo $\hat{\theta}_1 \leq \theta \leq \hat{\theta}_2$, a partir da amostra, com determinada probabilidade de conter o verdadeiro parâmetro populacional θ .

b) Testes de Hipóteses: a partir dos resultados extraídos da amostra selecionada aleatoriamente, testa-se os valores de certos parâmetros da população ou a natureza da sua distribuição.

Paramétricos: quando as hipóteses são formuladas com relação ao valor de um parâmetro populacional.

Aderência: quando as hipóteses são formuladas com relação à natureza da distribuição da população.

Teste de hipótese é uma regra de decisão, com base nos elementos amostrais, para aceitar ou rejeitar uma hipótese estatística.

Sendo as hipóteses estatísticas a serem testadas H_0 (hipótese nula) e H_1 (hipótese alternativa), com os exemplos:

Teste bicaudal (bilateral):

$$H_0 : \mu = 50Kg$$

$$H_1 : \mu \neq 50Kg$$

Teste unicaudal (unilateral) à direita:

$$H_0 : \mu = 50Kg$$

$$H_1 : \mu > 50Kg$$

Teste unicaudal (unilateral) à esquerda:

$$H_0 : \mu = 50Kg$$

$$H_1 : \mu < 50Kg$$

Erro tipo I ou erro de primeira espécie: comete-se ao rejeitar a hipótese H_0 quando ela é verdadeira, isto é, aceitar como diferentes médias de tratamentos que na verdade são iguais.

Erro tipo II ou erro de segunda espécie: comete-se ao não rejeitar a hipótese H_0 quando ela é falsa, isto é, aceitar como iguais médias de tratamentos que são diferentes.

Situação	Decisão	
	Não rejeitar H_0	Rejeitar H_0
H_0 : Verdadeira	Decisão correta	Erro tipo I
H_0 : Falsa	Erro tipo II	Decisão correta

O procedimento de erradamente rejeitar H_0 é considerado mais sério do que erroneamente não rejeitar H_0 .

1.2.11. Teste do valor médio versus um valor padrão populacional

$$H_0 : \mu = \mu_0 \quad H_1 : \begin{cases} \mu \neq \mu_0 \rightarrow \text{caso1} \\ \mu > \mu_0 \rightarrow \text{caso2} \\ \mu < \mu_0 \rightarrow \text{caso3} \end{cases}$$

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Caso 1: se $(-t_{\frac{\alpha}{2}} \leq t \leq t_{\frac{\alpha}{2}})$ não rejeitar H_0 ;

Caso 2: se $t > t_{\alpha}$ rejeitar H_0 ;

Caso 3: se $t < t_{\alpha}$ rejeitar H_0 .

Exercício 1.2.11.1 Uma determinada espécie florestal, na idade de sete anos, em sua região de origem possui uma produção padrão de $\mu_0 = 19 \text{ m}^3$ por 0,2 ha. Estimar o seu comportamento produtivo quando introduzida em outra região.

Tabela 1.2.11.1 - Volume de madeira ($\text{m}^3/\text{parcela}$).

18	25	19	21	25
23	17	22	15	27

a) Valor médio:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{212}{10} = 21,2 \text{ m}^3/0,2 \text{ ha}$$

b) Variância:

$$s^2 = \frac{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}{n-1} = \frac{4.632,0 - \frac{212^2}{10}}{9} = 15,2889$$

c) Desvio padrão:

$$s = \sqrt{s^2} = 3,91$$

d) Cálculo da estatística “t” de Student:

$$t = \frac{\frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}}{\frac{s}{\sqrt{n}}} = \frac{21,2 - 19,0}{\frac{3,91}{\sqrt{10}}} = 1,78 \text{ ns}$$

Em que: ns (não significante).

De acordo com a tabela A2, distribuição “t” de Student, tem-se: $t_{9; 0,05} = 2,26$.

Há razão para afirmar que a espécie florestal, em termos de produção de madeira, ao nível de significância $\alpha = 0,05$, apresentou o mesmo comportamento correspondente à região de origem.

1.2.12. Comparação dos valores médios para duas populações independentes

$$H_0 : \mu_1 = \mu_2 \text{ vs } H_1 : \mu_1 \neq \mu_2$$

Calcular:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2 \left(\frac{n_1 + n_2}{n_1 n_2} \right)}}$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Se $(-t_{\frac{\alpha}{2}} \leq t \leq t_{\frac{\alpha}{2}})$ não rejeitar H_0 .

Exercício 1.2.12.1. Comparar o comportamento produtivo de duas espécies florestais E1 e E2 com idades de 20 anos.

Tabela 1.2.12.1. Volume de madeira (parcelas de 0,2 ha).

Espécie 1		Espécie 2	
15	11	25	29
13	10	28	23
11	17	28	25
16	18	33	31
14		30	32

$$H_0 : \mu_1 = \mu_2 \text{ vs } H_1 : \mu_1 \neq \mu_2$$

a) Valores médios:

$$\bar{X}_1 = \frac{\sum_{i=1}^{n_1} X_{1i}}{n_1} = \frac{125}{9} = 13,89 \text{ m}^3 / 0,2 \text{ ha}$$

$$\bar{X}_2 = \frac{\sum_{i=1}^{n_2} X_{2i}}{n_2} = \frac{284}{10} = 28,40 \text{ m}^3 / 0,2 \text{ ha}$$

b) Variâncias:

$$s_1^2 = \frac{\sum_{i=1}^{n_1} X_{1i}^2 - \frac{(\sum_{i=1}^{n_1} X_{1i})^2}{n_1}}{n_1 - 1} = 8,1111$$

$$s_2^2 = \frac{\sum_{i=1}^{n_2} X_{2i}^2 - \frac{(\sum_{i=1}^{n_2} X_{2i})^2}{n_2}}{n_2 - 1} = 10,7111$$

c) Cálculo da estimativa da variância comum s^2 :

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)} = \frac{8 \times 8,1111 + 9 \times 10,7111}{17} = 9,4876$$

d) Cálculo da estatística “t”:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{28,40 - 13,89}{\sqrt{9,4876 \left(\frac{1}{9} + \frac{1}{10} \right)}} = 10,25^{**} \quad t_{17;0,01} = 2,8982$$

Em que: ** = altamente significativa ($\alpha = 0,01$).

Há razão para afirmar que as espécies florestais, ao nível de significância $\alpha = 0,01$, em termos de produção de madeira apresentam comportamentos diferentes. A espécie1 possui uma produção altamente significativa, isto é, diferente em relação à espécie2.

1.2.13. Comparação dos valores médios para duas populações dependentes. Teste “t” pareado

$$t = \frac{\hat{y} - 0}{\sqrt{\hat{V}(\hat{y})}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n} - 2 \frac{\text{Cov}(X_1, X_2)}{n}}} = \frac{\bar{d}}{s_{\bar{d}}} = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$$

$$\hat{y} = \bar{d} = \bar{X}_1 - \bar{X}_2$$

Exercício 1.2.13.1. Comparar o comportamento produtivo (volume de madeira) entre duas espécies consorciadas.

Tabela 1.2.13.1. Volume de E1 e E2 para nove parcelas de 0,5 ha.

E1	E2		E1	E2	
X_{1i}	X_{2i}	$d_i = X_{2i} - X_{1i}$	X_{1i}	X_{2i}	$d_i = X_{2i} - X_{1i}$
15	25	10	10	25	15
22	22	0	11	24	13
17	27	10	14	22	8
18	30	12	13	18	5
12	18	6			

a) Cálculo dos valores médios:

$$\bar{X}_1 = \frac{\sum_{i=1}^n X_{1i}}{n} = \frac{132}{9} = 14,67 \text{ m}^3/0,5ha \quad \bar{X}_2 = \frac{\sum_{i=1}^n X_{2i}}{n} = \frac{211}{9} = 23,44 \text{ m}^3/0,5ha$$

b) Cálculo das variâncias:

$$s_1^2 = \frac{\sum_{i=1}^n X_{1i}^2 - \frac{(\sum_{i=1}^n X_{1i})^2}{n}}{n-1} = 14,5000 \quad s_2^2 = \frac{\sum_{i=1}^n X_{2i}^2 - \frac{(\sum_{i=1}^n X_{2i})^2}{n}}{n-1} = 15,55278$$

c) Cálculo da covariância $Côv(X_1, X_2)$:

$$Côv(X_1, X_2) = \frac{\sum_{i=1}^n X_{1i} X_{2i} - \frac{(\sum_{i=1}^n X_{1i})(\sum_{i=1}^n X_{2i})}{n}}{n-1} = \frac{3.130,0 - \frac{132 \times 211}{9}}{8} = 4,4167$$

d) Cálculo da estatística “t”:

$$t = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{n} - 2Côv(X_1, X_2)}} = \frac{23,44 - 14,67}{\sqrt{\frac{14,5000}{9} + \frac{15,5278}{9} - 2 \times \frac{4,4167}{9}}} = 5,71^{**}$$

$$t_{8;0,01} = 3,3554$$

Há razão para afirmar que as espécies florestais, ao nível de significância $\alpha = 0,01$, em termos de produção de madeira, apresentam comportamentos diferentes. A espécie2 possui uma produção altamente significativa em relação à espécie1.

1.3. EXIGÊNCIAS DO MODELO MATEMÁTICO NA APLICAÇÃO DA ANÁLISE DE VARIÂNCIA

São as seguintes as suposições nos ensaios para a aplicação da análise de variância:

- Os efeitos devem ser aditivos. Os diversos efeitos devem ser aditivos e explicados por meio de um modelo matemático denominado linear aditivo;
- Os erros das observações devem ser independentes. A independência é assegurada pela aleatoriedade;
- Os erros das observações devem ser normalmente distribuídos;
- As variâncias das diferentes amostras devem ser homogêneas.

1.3.1. Teste de homogeneidade de variâncias pela razão máxima de Hartley

$$H = \frac{s_{Max}^2}{s_{Min}^2}$$

$$s_{Max}^2 : \text{maior variância} \quad s_{Min}^2 : \text{menor variância}$$

Compara o valor H com os valores críticos de $H_{(I,J-1)\alpha}$ da tabela de Pearson e Hartley (1970), tabela A9 do apêndice A, com I e $J-1$ graus de liberdade, em que I é o número de variâncias e $J-1$ é o número de graus de liberdade associado a cada variância. Se os tratamentos não possuem o mesmo número de repetições, utiliza-se a média deles.

Para $H \geq H_{(I,J-1)\alpha}$, rejeita-se a hipótese de homoscedasticidade ou homocedasticidade e se conclui pela não existência de homogeneidade de variâncias entre os tratamentos. Esse teste é eficiente para $I \leq 12$, quando todos os tratamentos tiverem o mesmo número de repetições. Um exemplo de aplicação está mostrado no exercício 2.5.1 do capítulo 2, considerando um experimento inteiramente casualizado com cinco tratamentos e seis repetições.

1.3.2. Homogeneidade de variâncias pelo teste de Bartlett

De acordo com Steel e Torrie (1960), o teste de Bartlett objetiva testar a homogeneidade de variâncias:

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2 \\ H_A : \text{Não } H_0 \end{cases}$$

O teste de Bartlett é dado pela estatística: $\chi_{I-1}^2 = \frac{M}{C}$.

$$M = \left[\sum_{i=1}^I (J_i - 1) \right] \ln \bar{\sigma}^2 - \sum_{i=1}^I (J_i - 1) \ln \sigma_i^2$$

$$M = 2,3026 \left\{ \left[\sum_{i=1}^I (J_i - 1) \right] \log \bar{\sigma}^2 - \sum_{i=1}^I (J_i - 1) \log \sigma_i^2 \right\}$$

$$C = 1 + \frac{1}{3(I-1)} \left[\sum_{i=1}^I \frac{1}{J_i - 1} - \frac{1}{\sum_{i=1}^I (J_i - 1)} \right]$$

$\bar{\sigma}^2$: é a média ponderada das estimativas de variâncias;

I : é o número de variâncias;

J_i : é o número de repetições do i -ésimo tratamento;

$\log$: logaritmo na base 10 e $\ln$: logaritmo neperiano.

No caso das estimativas de variâncias apresentarem o mesmo número de graus de liberdade:

$$M = 2,3026 \left[I(J-1) \log \bar{\sigma}^2 - (J-1) \sum_{i=1}^I \log \hat{\sigma}_i^2 \right]$$

$$C = 1 + \frac{1}{3(I-1)} \left[\frac{I}{J-1} - \frac{1}{I(J-1)} \right]$$

$$C = 1 + \frac{1}{3(I-1)(J-1)} \left(I - \frac{1}{I} \right) = 1 + \frac{1}{3(I-1)(J-1)} \left(\frac{I^2 - 1}{I} \right)$$

$$C = 1 + \frac{(I+1)(I-1)}{3I(I-1)(J-1)} = 1 + \frac{I+1}{3I(J-1)}$$

Um exemplo de aplicação está apresentado no exercício 2.5.1 do capítulo 2.

1.3.3. Homogeneidade de variâncias pelo teste de Cochran

$$C = \frac{\hat{\sigma}_{\max}^2}{\sum_{i=1}^I \hat{\sigma}_i^2}$$

A tabela do teste de Cochran, A7 do apêndice A, considera I e $J-1$ graus de liberdade, tal que I é o número de estimativas de variância e $(J-1)$ é o número de graus de liberdade associado a essas estimativas.

Um exemplo de aplicação está apresentado no exercício 2.5.1 do capítulo 2.

1.3.4. Teste de normalidade pelo Qui-Quadrado χ^2

A tabela 1.3.4.1 apresenta uma amostra com a ocorrência de k eventos possíveis $E_1, E_2, \dots, E_k$, os quais acontecem com as frequências observadas $o_1, o_2, \dots, o_k$. Por outro lado, de acordo com a natureza da pesquisa experimental utilizada, é esperado que tais frequências atendam a uma determinada regra de probabilidade e apresentem as frequências esperadas $e_1, e_2, \dots, e_k$. Deseja-se verificar se as frequências observadas diferem significativamente das esperadas (teóricas).

Tabela 1.3.4.1. Eventos: frequências observadas e esperadas.

Eventos	E_1	E_2	E_3	...	E_k
Freq. Observada	o_1	o_2	o_3	...	o_k
Freq. Esperada	e_1	e_2	e_3	...	e_k

A discrepância entre as frequências observadas e as esperadas pode ser avaliada pela estatística qui-quadrado (χ^2):

$$\chi_v^2 = \frac{(o_1 - e_1)^2}{e_1} + \frac{(o_2 - e_2)^2}{e_2} + \dots + \frac{(o_k - e_k)^2}{e_k} = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \quad (1.3.4.1)$$

$$\sum_{i=1}^k o_i = \sum_{i=1}^k e_i = N$$

Pode ser observado que quanto maior for o valor de χ^2 maior será a discrepância entre as frequências observadas e as esperadas.

A estatística 1.3.4.1 pode ser usada para testar quão, aproximadamente, as distribuições teóricas, tais como a Normal, Binomial, Poisson etc, ajustam-se às distribuições empíricas, ou seja, as calculadas a partir dos dados amostrais.

O número de graus de liberdade (ν) é dado por:

a) $\nu = k - 1$, quando as frequências esperadas forem calculadas sem fazer estimativas de parâmetros populacionais;

b) $\nu = k - 1 - p$, quando as frequências esperadas são obtidas a partir de estimativas de p parâmetros populacionais.

Exercício 1.3.4.1. Em uma amostra aleatória com 105 mudas, coletadas em um viveiro florestal, em que foi medida a altura total (em cm) de cada planta, testar a aderência das observações da variável altura total à distribuição normal pelo teste de Qui-Quadrado (χ^2).

Tabela 1.3.4.2. Distribuição das 105 árvores ($k = 6$ classes).

Altura	F_i	X_i	Limites L_i	Z_i	Áreas	Prob.	Freq. Esperada
15-17	6	16	15	-2,49	0,4936	0,0391	4,11
17-19	14	18	17	-1,69	0,4545	0,1439	15,11
19-21	25	20	19	-0,88	0,3106	0,2787	29,26
21-23	38	22	21	-0,08	0,0319	0,2961	31,09
23-25	15	24	23	0,72	0,2642	0,1728	18,14
25-27	7	26	25	1,53	0,4370	0,0531	5,58
			27	2,33	0,4901		
Total	105						

H_0 : A variável resposta altura de mudas pode ser estudada pela distribuição normal; H_A : Não H_0 .

$$\bar{X} = \frac{\sum_{i=1}^k F_i X_i}{\sum_{i=1}^k F_i} = \frac{2.226,0}{105} = 21,2 \text{ cm}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^k F_i X_i^2 - \frac{(\sum_{i=1}^k F_i X_i)^2}{\sum_{i=1}^k F_i}}{\sum_{i=1}^k F_i - 1} = \frac{47.836,0 - \frac{(2.226,0)^2}{105}}{104} = \frac{644,8}{104} = 6,2 \quad \hat{\sigma} = 2,49$$

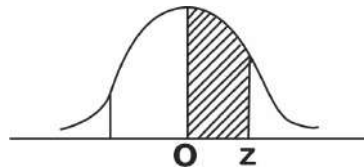
$$Z_i = \frac{L_i - \bar{X}}{\hat{\sigma}}$$

$$Z_1 = \frac{15 - 21,2}{2,49} = -2,49 \quad Z_2 = \frac{17 - 21,2}{2,49} = -1,69 \quad Z_3 = \frac{19 - 21,2}{2,49} = -0,88$$

$$Z_4 = \frac{21 - 21,2}{2,49} = -0,08 \quad Z_5 = \frac{23 - 21,2}{2,49} = 0,72 \quad Z_6 = \frac{25 - 21,2}{2,49} = 1,53$$

$$Z_7 = \frac{27 - 21,2}{2,49} = 2,33$$

Na coluna “Áreas”, os valores foram obtidos na tabela A1 do Apêndice A, que apresenta os limites da distribuição normal reduzida.



$$\int_z^0 \frac{1}{\sqrt{2\pi}} \text{Exp}\left(-\frac{1}{2} Z^2\right)$$

$$\chi^2_{k-p-1} = \frac{(6-4,11)^2}{4,11} + \frac{(14-15,11)^2}{15,11} + \frac{(25-29,26)^2}{29,26} + \frac{(38-31,09)^2}{31,09} + \frac{(15-18,14)^2}{18,14} + \frac{(7-5,58)^2}{5,58}$$

$$\chi^2_{6-2-1} = 0,86912 + 0,08154 + 0,62022 + 1,53580 + 0,54353 + 0,36136 = 4,012$$

Em que $p = 2$ é o número de parâmetros estimados ($\hat{m}$ e $\hat{\sigma}$).

$$\chi^2_{CAL.} = 4,012 \quad \chi^2_{3;0,05} = 7,81 \quad \chi^2_{CAL.} < \chi^2_{3;0,05}$$

Não há razão para rejeitar H_0 , isto é, a variável altura das plantas do viveiro pode ser estudada usando a distribuição normal $N(\mu; \sigma^2) = N(21,2 \text{ cm}; 6,2 \text{ cm}^2)$.

Existem outros testes para verificar a suposição de normalidade dos dados, tais como: Kolmogorov-Smirnov, Cramér-Von Mises, Anderson-Darling e Shapiro-Wilk, bem como recursos gráficos, como histograma e normal plot.

As estatísticas desses testes possuem metodologias diferentes para realização do teste de hipóteses. Os testes de Anderson-Darling, Cramér-VonMises e Kolmogorov-Smirnov são baseados na função de distribuição empírica (FDE) dos dados e apresentam vantagens sobre o teste de aderência Qui-Quadrado, incluindo maior poder e invariância em relação aos pontos médios dos intervalos escolhidos. O teste de Kolmogorov-Smirnov considera a maior diferença entre a distribuição empírica e a hipotética. Os testes Anderson-Darling e Cramér-VonMises pertencem à classe quadrática de estatísticas baseadas na FDE, pois operam com as diferenças quadráticas entre a distribuição empírica e a hipotética. O teste de Shapiro-Wilk está baseado nos valores amostrais ordenados elevados ao quadrado.

É importante esclarecer, no caso da verificação de normalidade de um experimento, que a exigência de normalidade não deve ser feita para os dados observados nas unidades experimentais, mas para a distribuição dos resíduos do modelo linear utilizado. O programa SAS, **Prog5a**, em anexo no

apêndice B, faz a verificação de normalidade para um ensaio casualizado em blocos, exercício 4.2.1 do capítulo 4, abrangendo os testes Kolmogorov-Smirnov, Cramér-Von Mises, Anderson-Darling e Shapiro-Wilk.

1.3.5. Teste de aderência de Kolmogorov-Smirnov (KS)

O teste de Kolmogorov-Smirnov (KS), Kolmogorov (1933), é um método não paramétrico de aderência que compara a distribuição acumulada de uma distribuição específica e conhecida $F(X)$, por exemplo, a distribuição normal de dados provenientes de uma distribuição desconhecida $F_0(X)$. Determina se os valores de uma amostra $X_1, X_2, \dots, X_n$ podem ser considerados como provenientes de uma população com a distribuição teórica $F(X)$. A hipótese de nulidade especifica a distribuição $F(X)$.

De acordo com Campos (1979), se uma amostra aleatória $X_1, X_2, \dots, X_n$, denotando as estatísticas de ordem sem observações repetidas, é retirada de alguma população cuja distribuição $F_0(X)$ é desconhecida, estabelece-se o confronto com $F(X)$ para verificar se é razoável estudar os dados por meio de $F(X)$, a qual é admitida como verdadeira função de distribuição da amostra casualizada.

Uma vantagem do KS em relação ao teste de aderência Qui-Quadrado (χ^2) é ser aplicado sem restrição para pequenas amostras, assim como opera com os dados individualmente, isto é, não perde informações devido à construção de agrupamentos. No caso de pequenas amostras, é mais poderoso em comparação com o teste de χ^2 .

O teste de Komogorov e Smirnov (KS) é mais confiável quando a comparação é feita com uma distribuição normal conhecida, isto é, quando os valores da média e da variância são conhecidos. Por outro lado, caso não seja possível aplicar o KS, então existe o teste de Lilliefors, o qual é uma modificação do KS e apresenta interpretações similares, sendo confiável para estimar a curva normal a partir dos dados da amostra.

O teste Komogorov-Smirnov (KS) se baseia na distância máxima entre uma distribuição teórica de referência conhecida $F(X)$ e uma distribuição que foi observada empiricamente $F_0(X)$.

Dadas as seguintes definições:

$F(X)$: é a proporção de valores esperados $\leq X$;

$S(X)$: é a proporção de valores observados $\leq X$. É uma função empírica de distribuição e um estimador de $F(X)$;

$F_0(X)$: é a verdadeira função de distribuição de X , a qual é desconhecida.

Teste de hipótese bilateral:

$H_0 : F = F_0$: para todo $-\infty < X < \infty$;

$H_1 : F \neq F_0$: para pelo menos um valor de X .

Como a função empírica $S(X)$ é descontínua e a função de distribuição hipotética é contínua, então para a realização do teste KS , têm-se as seguintes alternativas:

$$D_1 = \sup_x |F(X_i) - S(X_i)| \quad D_2 = \sup_x |F(X_i) - S(X_{i-1})|$$

As estatísticas $D_1 = |F(X_i) - S(X_i)|$ e $D_2 = |F(X_i) - S(X_{i-1})|$ medem as diferenças (verticais) entre os gráficos das duas funções acumuladas, isto é, teórica e empírica nos pontos $X_{(i-1)}$ e X_i . Para a determinação dos supremos, em cada ponto X_i , considera-se a maior das diferenças. Assim, pode-se utilizar a seguinte estatística para o teste:

$$D = \max(D_1, D_2)$$

Rejeita-se H_0 para a normalidade dos dados, com $\alpha - 1$ de confiança, se $D \geq d$. Caso contrário não se rejeita a hipótese de normalidade. Os valores de (d) são encontrados na tabela A6 do apêndice A, em que $P_0(D \geq d) = \alpha$.

De acordo com Pimentel Gomes e Nogueira (1980), sendo C um conjunto de números reais. O supremo desse conjunto é, por definição, o número real que tenha as seguintes propriedades:

- Nenhum número do conjunto C supera M ;
- Dado um número real qualquer $M' < M$, portanto, existe sempre pelo menos um número de C maior que M' .

Quando o supremo pertence ao conjunto C , então ele é dito máximo de C .

- O teste KS pode, sem restrição, ser aplicado em pequenas amostras;
- A distribuição da estatística KS independe da função de distribuição acumulada usada no teste;
- Para garantir que a região crítica do teste seja válida, então a distribuição a ser testada deve ser completamente especificada;
- O teste é mais sensível próximo ao centro da distribuição do que nas caudas.

Exercício 1.3.5.1. Em uma amostra de 15 observações referentes à altura (em cm) de mudas de uma espécie florestal, testar, pelo método de Kolmogorov-Smirnov (KS), se a variável altura possui uma distribuição normal com média $\mu = 13,3\text{cm}$ e variância $\sigma^2 = 1,7^2 = 2,89 \text{ cm}^2$.

Alturas das plantas				
13,7	11,0	13,8	14,8	14,2
15,4	13,3	16,2	15,8	16,0
14,1	13,6	14,7	12,1	11,3

H_0 : os dados seguem uma distribuição de probabilidade normal;

H_1 : os dados não seguem uma distribuição de probabilidade normal.

Na tabela 1.3.5.1, a coluna X apresenta os valores da amostra aleatória ordenados na forma de estatística de ordem. A coluna $S(X_i)$ apresenta as proporções, considerando a posição da observação em relação ao tamanho da amostra, isto é, K/n . No presente exemplo, como são 15 observações, a proporção do primeiro elemento será $1/15 = 0,06667$, o segundo $2/15 = 0,13333$, e assim por diante. O último valor corresponderá a um (1). Por outro lado, quando o valor de $S(X_i)$ é considerado no passo anterior, então o primeiro elemento receberá o valor zero (0), ou seja, $|F(X_1) - 0| = 0,088037$.

Na construção do teste é necessário calcular as probabilidades acumuladas da função teórica. Nesse exemplo da distribuição normal a coluna $F(X_i)$ mostra os valores das probabilidades acumuladas da função teórica $F(X)$.

$$F(X_i) = \int_{-\infty}^{X_i} \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{X_i - \mu}{\sigma}\right)^2\right] d_x \quad F(X_i) = \int_{-\infty}^{X_i} \frac{1}{1,7\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{X_i - 13,3}{1,7}\right)^2\right] d_x$$

Tabela 1.3.5.1. Teste de aderência Kolmogorov-Smirnov (KS).

Arv.	X_i	$F(X_i)$	$S(X_i)$	$ F(X_i) - S(X_i) $	$ F(X_i) - S(X_{i-1}) $
1	11,0	0,088037	0,06667	0,021367	$ F(X_1) - 0 = 0,088037$
2	11,3	0,119703	0,13333	0,013627	$ F(X_2) - S(X_1) = 0,053033$
3	12,1	0,240131	0,20000	0,040131	$ F(X_3) - S(X_2) = 0,106801$
4	13,3	0,500000	0,26667	0,233330	$ F(X_4) - S(X_3) = 0,300000$
5	13,6	0,570038	0,33333	0,236708	$ F(X_5) - S(X_4) = 0,303368^*$
6	13,7	0,593010	0,40000	0,193010	$ F(X_6) - S(X_5) = 0,259680$
7	13,8	0,615666	0,46667	0,148996	$ F(X_7) - S(X_6) = 0,215666$
8	14,1	0,681033	0,53333	0,147703	$ F(X_8) - S(X_7) = 0,214363$
9	14,2	0,701740	0,60000	0,101740	$ F(X_9) - S(X_8) = 0,168410$
10	14,7	0,794896	0,66667	0,128226	$ F(X_{10}) - S(X_9) = 0,194896$
11	14,8	0,811207	0,73333	0,077877	$ F(X_{11}) - S(X_{10}) = 0,144537$
12	15,4	0,891639	0,80000	0,091639	$ F(X_{12}) - S(X_{11}) = 0,158309$
13	15,8	0,929299	0,86667	0,062629	$ F(X_{13}) - S(X_{12}) = 0,129299$
14	16,0	0,943883	0,93333	0,010553	$ F(X_{14}) - S(X_{13}) = 0,077213$
15	16,2	0,955985	1,00000	0,044015	$ F(X_{15}) - S(X_{14}) = 0,022655$

Os valores da coluna $F(X_i)$ podem ser obtidos nas tabelas de distribuição normal reduzida. Nesse exemplo, os valores de $F(X_i)$ foram calculados pelo programa SAS, **Prog1**, em anexo no apêndice B.

Como $n = 15$ e $\alpha = 0,05$, (tabela A6 do apêndice A), como: $P_0(D \geq 0,338) = 0,05$, pois $|F(X_5) - S(X_4)| = 0,303$, há razão para admitir que as observações (alturas) são normalmente distribuídas, isto é: $X_i \sim N(\mu = 13,3 \text{ cm}, \sigma^2 = 2,89 \text{ cm}^2)$. Não se rejeita H_0 :

Rejeita-se : $H_0 : F = F_0$ vs $H_1 : F \neq F_0$, (para pelo menos um valor de X , se $D \geq d$).

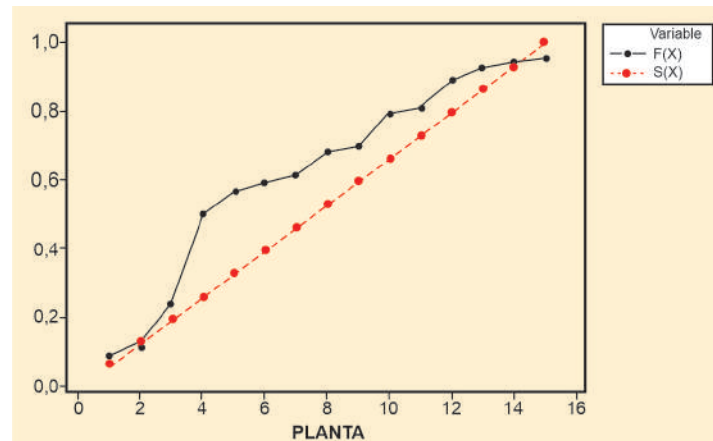


Figura 1.3.5.1. Gráfico das duas funções acumuladas $F(X)$ e $S(X)$.

A Figura 1.3.5.1 mostra o gráfico, considerando as duas funções acumuladas, isto é, a teórica $F(X)$ e a empírica $S(X)$.

1.3.6. Teste de aderência de Lilliefors

Diferentemente do teste de Kolmogorov-Smirnov (KS), que adota uma distribuição de probabilidade determinada, isto é, com média e variâncias conhecidas (especificadas), o teste de Lilliefors (1967) efetua uma modificação no teste KS , em que, ao invés de especificadas, a média e a variância são estimadas por meio dos dados amostrados.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}{n-1}$$

De acordo com Campos (1979) — considerando que, por meio da amostra, a média e a variância são estimadas, ou seja, não são especificadas —, os limites superiores da distribuição de D (bilateral) do teste de Kolmogorov-Smirnov não são apropriados. Assim sendo, limites superiores foram criados empiricamente (tabela A11 do apêndice A) para o teste de Lilliefors.

Exercício 1.3.6.1. Em uma amostra de 15 observações, citada no exercício 1.3.5.1, referente à altura (em cm) de mudas de uma espécie florestal, testar, pelo método de Lilliefors, se a variável altura possui uma distribuição normal com média $\mu = 14$ cm e variância $\sigma^2 = 2,5640$ cm². Tais valores correspondem aos estimados por meio dos dados amostrais.

H_0 : os dados seguem uma distribuição de probabilidade normal;

H_1 : os dados não seguem uma distribuição de probabilidade normal.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = 14 \text{ cm} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}}{n-1} = 2,5640 \text{ cm}^2$$

A coluna $F(X_i)$, tabela 1.3.6.1, mostra os valores das probabilidades acumuladas da função teórica $F(X)$ considerada normal com $\bar{X} = 14 \text{ cm}$ e $\hat{\sigma}^2 = 2,5640 \text{ cm}^2$. Os valores da coluna $F(X_i)$ podem ser obtidos em tabelas de distribuição normal reduzida. Nesse exemplo, os valores de $F(X_i)$ foram calculados pelo programa SAS, **Prog2**, em anexo no apêndice B.

Tabela 1.3.6.1. Dados e cálculos requeridos para o teste de Lilliefors.

Arv.	X_i	Z_i	$F(X_i)$	$S(X_i)$	$ F(X_i) - S(X_i) $	$ F(X_i) - S(X_{i-1}) $
1	11,0	-1,87343	0,030497	0,06667	0,036173	$ F(X_1) - 0 = 0,036173$
2	11,3	-1,68609	0,045880	0,13333	0,087450	$ F(X_2) - S(X_1) = 0,020790$
3	12,1	-1,18651	0,117698	0,20000	0,082302	$ F(X_3) - S(X_2) = 0,015632$
4	13,3	-0,43713	0,330998	0,26667	0,064328	$ F(X_4) - S(X_3) = 0,130998$
5	13,6	-0,24979	0,401369	0,33333	0,068039	$ F(X_5) - S(X_4) = 0,134699^*$
6	13,7	-0,18734	0,425692	0,40000	0,025692	$ F(X_6) - S(X_5) = 0,092362$
7	13,8	-0,12490	0,450300	0,46667	0,016370	$ F(X_7) - S(X_6) = 0,050300$
8	14,1	0,06245	0,524898	0,53333	0,008432	$ F(X_8) - S(X_7) = 0,058228$
9	14,2	0,12490	0,549700	0,60000	0,050300	$ F(X_9) - S(X_8) = 0,016370$
10	14,7	0,43713	0,669002	0,66667	0,002332	$ F(X_{10}) - S(X_9) = 0,069002$
11	14,8	0,49958	0,691325	0,73333	0,042005	$ F(X_{11}) - S(X_{10}) = 0,024655$
12	15,4	0,87427	0,809027	0,80000	0,009027	$ F(X_{12}) - S(X_{11}) = 0,075697$
13	15,8	1,12406	0,869519	0,86667	0,002849	$ F(X_{13}) - S(X_{12}) = 0,069519$
14	16,0	1,24895	0,894172	0,93333	0,039158	$ F(X_{14}) - S(X_{13}) = 0,027502$
15	16,2	1,37385	0,915268	1,00000	0,084732	$ F(X_{15}) - S(X_{14}) = 0,018062$

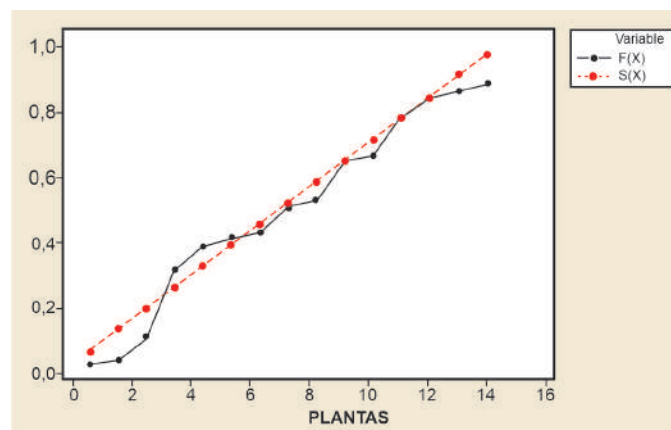


Figura 1.3.6.1. Gráfico das duas funções acumuladas $F(X)$ e $S(X)$.

A Figura 1.3.6.1 mostra o gráfico considerando as duas funções acumuladas, isto é, a teórica $F(X)$ e a empírica $S(X)$.

$H_0: F = F_0$ vs $H_1: F \neq F_0$, para pelo menos um valor de X , rejeita-se H_0 se $D \geq d$.

Considerando $n = 15$ e $\alpha = 0,05$, tabela A7, tem-se: $P_0(D \geq 0,220) = 0,05$. Como $|F(X_5) - S(X_4)| = 0,135$, então não se rejeita H_0 , ou seja, há razão para admitir pelo teste de Lilliefors que as observações (alturas) são normalmente distribuídas, isto é:

$$X_i \sim N(\mu=14 \text{ cm}, \sigma^2 = 2,5640 \text{ cm}^2)$$

1.3.7. Teste de aderência de Cramér-Von Mises

O teste de Cramér-Von Mises apresenta as mesmas finalidades do teste bilateral de Kolmogorov-Smirnov. Dada uma amostra aleatória $X_1, X_2, \dots, X_n$ — denotando as estatísticas de ordem, então: sem observações repetidas —, a qual foi retirada de alguma população, cuja distribuição $F_0(X)$ é desconhecida, então se estabelece o confronto com $F(X)$.

Seja a hipótese: $H_0 : F = F_0$ vs $H_0 : F \neq F_0$ (para algum valor de X).

De acordo com Campos (1979), tem-se:

$$D_1 = \frac{1}{12n} + \sum_{i=1}^n \left\{ F[X_{(i)}] - \frac{2i-1}{2n} \right\}^2$$

Rejeita-se $H_0 : F = F_0$, ao nível α de significância, se $D_1 \geq d_\alpha$, tal que:

$$P_0[D_1 \geq d_\alpha] = \alpha$$

Os limites aproximados, da distribuição de D_1 , são:

$d_{(0,20)} = 0,241$	$d_{(0,10)} = 0,347$	$d_{(0,05)} = 0,461$	$d_{(0,01)} = 0,743$
----------------------	----------------------	----------------------	----------------------

Exercício 1.3.7.1. Conhecendo os dados já ordenados na forma crescente de uma amostra de cinco alturas de árvores, verificar se a variável altura, pelo teste de Cramér-Von Mises, segue a distribuição normal.

Os valores da coluna $F(X_i)$ podem ser obtidos em tabelas de distribuição normal reduzida. Nesse exemplo, os valores de $F(X_i)$ foram calculados pelo programa SAS, **Prog3**, em anexo no apêndice B.

Tabela 1.3.7.1. Teste de normalidade pelo teste Cramér-Von Mises.

(i)	$X_{(i)}$	$Z_{(i)}$	$F(X_{(i)})$	$C = \frac{2i-1}{2 \times 5}$	$[F(X_{(i)}) - C]^2$
1	11,8	-1,0830	0,139299	0,1	0,0015444
2	11,9	-0,9627	0,167752	0,3	0,0174895
3	12,8	0,1203	0,547911	0,5	0,0022955
4	13,4	0,8424	0,800302	0,7	0,0100605
5	13,6	1,0830	0,860701	0,9	0,0015444
					0,03293

H_0 : Os dados seguem uma distribuição de probabilidade normal;

H_1 : Os dados não seguem uma distribuição de probabilidade normal.

$$D_1 = \frac{1}{12n} + \sum_{i=1}^n \left[F(X_{(i)}) - \frac{2i-1}{2n} \right]^2 = \frac{1}{12 \times 5} + 0,03293 = 0,0496$$

Dado que $D_1 < 0,461$, considerando o nível de significância $\alpha = 0,05$, há razão para considerar que a variável altura siga a distribuição normal. Na Tabela 1.3.9.2, o teste de normalidade realizado no SAS apresenta o seguinte resultado:

Cramér-Von Mises	Q	0,049601	Pr > W-Sq	0,2500
------------------	---	----------	-----------	--------

1.3.8. Teste de aderência de Anderson-Darling

É uma alternativa aos testes de Qui-Quadrado e de Kolmogorov-Smirnov. A sua metodologia é próxima ao teste *KS*, o qual apresenta uma forma da função distância para calcular a similaridade entre $S(X)$ (distribuição empírica) e $F(X)$ (distribuição teórica). Porém, ao contrário do *KS*, a estatística *AD* considera o intervalo total de valores dos dados, ao invés de simplesmente o máximo desvio.

Dada uma amostra aleatória $X_1, X_2, \dots, X_n$ denotando as estatísticas de ordem. Então:

$$AD = \int_{-\infty}^{\infty} \frac{[S(X) - F(X)]^2}{F(X)[1 - F(X)]} dF(X)$$

Podendo ser usada a forma simplificada:

$$AD = -n - \frac{1}{n} \times \sum_{i=1}^n (2i-1) \{ \ln[F(X_i)] + \ln[1 - F(X_{(n+1-i)})] \}$$

$F(X_i)$: é a distribuição acumulada para a distribuição específica e (i) corresponde ao índice quando os dados estão ordenados na ordem crescente.

Quando se tratar de pequenas amostras, o valor de *AD* deve ser ajustado para:

$$AD^* = AD \left(1 + \frac{0,75}{n} + \frac{2,25}{n^2} \right)$$

E o *p*-valor é calculado a partir da estatística modificada AD^* , isto é:

- i) Se $AD^* \geq 0,6 \rightarrow p = \exp[1,2937 - 5,709(AD^*) + 0,0186(AD^*)^2]$;
- ii) Se $0,34 \leq AD^* < 0,6 \rightarrow p = \exp[0,9177 - 4,279(AD^*) - 1,38(AD^*)^2]$;
- iii) Se $0,2 \leq AD^* < 0,34 \rightarrow p = 1 - \exp[-8,318 + 42,796(AD^*) - 59,938(AD^*)^2]$;
- iv) Se $AD^* < 0,2 \rightarrow p = 1 - \exp[-13,436 + 101,14(AD^*) - 223,73(AD^*)^2]$.

Exercício 1.3.8.1. Sendo os dados ordenados na forma crescente de uma amostra simples de cinco alturas de árvores, testar as hipóteses pelo teste de Anderson-Darling.

H_0 : os dados seguem uma distribuição de probabilidade normal;

H_1 : os dados não seguem uma distribuição de probabilidade normal.

Tabela 1.3.8.1. Teste de normalidade de Anderson-Darling.

(i)	$X_{(i)}$	$Z_{(i)}$	$F(X_{(i)})$	$F(X_{(n+1-i)})$	$\ln[F(X_{(i)})]$	$\ln[1 - F(X_{(n+1-i)})]$
1	11,8	-1,0830	0,139299	0,860701	-1,97113	-1,97113
2	11,9	-0,9627	0,167752	0,800302	-1,78527	-1,61095
3	12,8	0,1203	0,547911	0,547911	-0,60164	-0,79388
4	13,4	0,8424	0,800302	0,167752	-0,22277	-0,18362
5	13,6	1,0830	0,860701	0,139299	-0,15001	-0,15001

Tabela 1.3.8.1. Teste de normalidade de Anderson-Darling (continuação).

$2i - 1$	$\ln[F(X_{(i)})]$	$\ln[1 - F(X_{(n+1-i)})]$	$(2i - 1) \times \{\ln[F(X_{(i)})] + \ln[1 - F(X_{(n+1-i)})]\}$
1	-1,97113	-1,97113	-3,9423
3	-1,78527	-1,61095	-10,1887
5	-0,60164	-0,79388	-6,9776
7	-0,22277	-0,18362	-2,8447
9	-0,15001	-0,15001	-2,7002
Total			-26,6535

Os valores da coluna $F(X_{(i)})$ podem ser obtidos em tabelas de distribuição normal reduzida. Nesse exemplo, os valores de $F(X_{(i)})$ foram calculados pelo programa SAS, **Prog3**, em anexo no apêndice B.

$$AD = -n - \frac{1}{n} \times \sum_{i=1}^n (2i - 1) \{ \ln[F(X_{(i)})] + \ln[1 - F(X_{(n+1-i)})] \}$$

$$AD = -5 - \frac{1}{5}(-26,6535) = 0,3307$$

A tabela 1.3.9.2 mostra o resultado do teste de Anderson Darling no SAS, em que o valor da estatística do teste 0,3307679 corresponde ao p -valor igual a 0,25. Então, conclui-se que os dados seguem uma distribuição de probabilidade normal.

Anderson-Darling	A-Sq	0,330679	Pr > A-Sq	0,2500
------------------	------	----------	-----------	--------

Mas, como se trata de uma pequena amostra, então o valor de AD deve ser ajustado:

$$AD^* = AD \left(1 + \frac{0,75}{n} + \frac{2,25}{n^2} \right) = 0,3307 \times \left(1 + \frac{0,75}{5} + \frac{2,25}{5^2} \right) = 0,410068$$

Como $0,34 \leq AD^* < 0,6$, então: $p = \exp[0,9177 - 4,279(AD^*) - 1,38(AD^*)^2] = 0,3433$

Dado que o p -valor é igual a 0,3433, há razão para considerar que a variável altura siga a distribuição normal. A figura 1.3.8.1 mostra o resultado obtido no Programa de Análise Estatística Minitab, o qual utiliza AD^* corrigido para pequenas amostras.

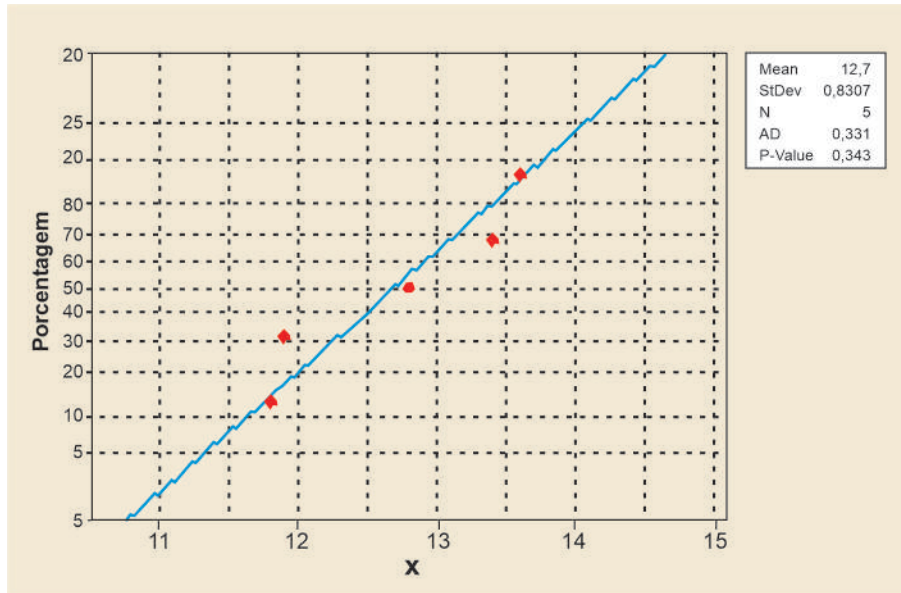


Figura 1.3.8.1. Teste de Anderson-Darling no Minitab.

1.3.9. Teste de Shapiro-Wilk (1965)

O teste de Shapiro-Wilk objetiva testar a hipótese nula para verificar se uma amostra de tamanho n , $X_1, X_2, \dots, X_n$, retirada de uma população, pode ser considerada como pertencente a uma distribuição normal.

Dada uma amostra aleatória já ordenada na forma $X_1 \leq X_2 \leq \dots \leq X_n$ (estatísticas de ordem), para obter o valor da estatística W_p de Shapiro-Wilk, deve-se proceder da seguinte forma:

a) Calcular a soma de quadrado dos desvios:

$$SQ_{DESVIOS} = \sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}$$

b) Considerar que (n) é o número de unidades de amostra. Desse modo:

b1) se (n) é um número par, então considerar $n = 2k$, em que $k = \frac{n}{2}$ e obter:

$$B = \sum_{i=1}^k a_{n-i+1} (X_{(n-i+1)} - X_i)$$

b2) se (n) é um número ímpar, então considerar $n = 2k + 1$, em que $k = \frac{n-1}{2}$. O cálculo é similar ao item anterior, haja vista que $a_{k+1} = 0$ quando $n = 2k + 1$. Assim sendo, determinar:

$$B = a_n(X_n - X_1) + \dots + a_{k+2}(X_{k+2} - X_k)$$

O valor X_{k+1} , que corresponde à mediana, não é considerado no cálculo.

No teste de Shapiro-Wilk, os valores dos coeficientes de (a_i) são tabelados, entretanto é preciso que os dados observados estejam ordenados na forma crescente. As colunas na tabela 1.3.9.1 representam o número de observações e as linhas representam a posição ordenada das observações.

Tabela 1.3.9.1. Valores dos coeficientes (a_i) do teste de Shapiro-Wilk.

(i)	Valores dos coeficientes do teste de Shapiro-Wilk								
	2	3	4	5	6	7	8	9	10
1	0,7071	0,7071	0,6872	0,6646	0,6431	0,6233	0,6052	0,5888	0,5739
2			0,1677	0,2413	0,2806	0,3031	0,3164	0,3244	0,3291
3					0,0875	0,1401	0,1743	0,1976	0,2141
4							0,0561	0,0947	0,1224
5									0,0399

c) Calcular a estatística:

$$W_C = \frac{B^2}{SQ_{DESVIOS}}$$

d) Comparar o valor W_C com o valor tabelar $W_{n;\alpha}$. Se $W_C < W_{n;\alpha}$, ao nível de significância α , não se rejeita a hipótese de normalidade. O teste de Shapiro-Wilk é:

- d1) específico para a distribuição normal e possui alto poder de decisão;
- d2) limitado para amostras com tamanho entre 3 e 50 observações;
- d3) sensível para valores iguais ou muito próximos.

Exercício 1.3.9.1. Sendo os dados já ordenados na forma crescente de uma amostra simples de cinco alturas de árvores, testar as hipóteses pelo teste de Shapiro-Wilk.

11,8	11,9	12,8	13,4	13,6
------	------	------	------	------

H_0 : os dados seguem uma distribuição de probabilidade normal;

H_1 : os dados não seguem uma distribuição de probabilidade normal.

$$SQ_{DESVIOS} = \sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n} = 809,21 - \frac{63,5^2}{5} = 2,76$$

Como $n = 5$ é um número ímpar, então $n = 2k + 1$, logo $k = 2$. Assim sendo:

$$B = a_n(X_n - X_1) + \dots + a_{k+2}(X_{k+2} - X_k)$$

$$B = a_5(X_5 - X_1) + a_4(X_4 - X_2)$$

Sejam os coeficientes $a_{(n-i-1)}$, tabela 1.3.9.1, do teste de Shapiro-Wilk: $a_5 = 0,6646$, $a_4 = 0,2413$ e $a_3 = 0$ ($a_{k+1} = 0$ quando $n = 2k + 1$), isto é $k = \frac{n-1}{2} = 2$, então:

$$B = 0,6646(13,6 - 11,8) + 0,2413(13,4 - 11,9) = 1,55823$$

Obter o valor da estatística W_C :

$$W_C = \frac{B^2}{SQ_{DESVIOS}} = \frac{1,55823^2}{2,76} = 0,879$$

Consultando a tabela A18 (anexo A) que apresenta os quantis de ordem ($p = 1 - \alpha$) para o teste de Shapiro-Wilk (W_p), verifica-se que o valor calculado (0,879) está entre as probabilidades 0,10 e 0,50. Logo não se rejeita, ao nível de 5 % de significância, a hipótese de normalidade dos erros.

n	$W_{0,01}$	$W_{0,02}$	$W_{0,05}$	$W_{0,10}$	$W_{0,50}$	$W_{0,90}$	$W_{0,95}$
5	0,686	0,715	0,762	0,803	0,927	0,979	0,986

A análise realizada no SAS, tabela 1.3.9.2, mostra que para $W_C = 0,879$ o p -valor é igual a 0,3090. Portanto, há razão para considerar que a variável altura, ao nível de $\alpha = 0,05$ de significância, siga a distribuição normal.

Tabela 1.3.9.2. Teste de normalidade para os quatro testes no SAS.

Testes de Normalidade				
Testes	Estatística		p-valor	
Shapiro-Wilk	W	0,879921	Pr < W	0,3090
Kolmogorov-Smirnov	D	0,232248	Pr > D	0,1500
Cramér-Von Mises	Q	0,049601	Pr > W-Sq	0,2500
Anderson-Darling	A-Sq	0,330679	Pr > A-Sq	0,2500

O SAS realiza o teste de normalidade para os dados amostrais. O “proc univariate normal” possibilita aplicar os seguintes testes: Kolmogorov-Smirnov, Cramér-Von Mises, Anderson-Darling e Shapiro-Wilk. Nesse exemplo, as análises foram realizadas pelo programa SAS **Prog4**, em anexo no apêndice B.

1.4. BIBLIOGRAFIA

- CAMPOS, H. **Estatística Experimental não paramétrica**. 3ª ed. Piracicaba: ESALQ-USP, 1979. 343 p.
- FISCHER, A. R. *The Design of Experiments*. Edinburgh and London: Oliver and Boyd, 1935.
- KOLMOGOROV, A. N. Sulla determinazione empirica di una legge distribuzione. **Giornale dell'Istituto Italiano degli Attuari**, n. 4, p. 83-91, 1933.
- LILLIEFORS, H. W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. **J. Amer. Statist. Ass.**, 62: 399 - 402, 1967.
- PADOVANI, C R. **Delineamento de experimentos**. São Paulo: Cultura Acadêmica: Universidade Estadual Paulista, Pró-Reitoria de Graduação, 2014. 128 p.
- PEARSON, E. S.; HARTLEY, H. O. **Biometrika tables for statisticians**. V. 1. Cambridge, University Press, 1970. 270 p.
- PERES, C. A.; SALDIVA, C. D. **Planejamento de experimentos**. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA. São Paulo, 1982. 98p.
- PIMENTEL GOMES, F.; **Estatística moderna na pesquisa agropecuária**. Piracicaba: POTAFOS, 1984. 160 p.
- PIMENTEL GOMES, F.; NOGUEIRA, I. R. **Análise matemática**. 2ª ed. Piracicaba: ESALQ-USP, 1980. 371 p.
- SALSBURG, D. **Uma senhora toma chá... como a estatística revolucionou a ciência no século XX**. Rio de Janeiro: Zahar, 2009. 286 p.
- SHAPIRO, S. S.; WILK, M.B. An Analysis of Variance Test for Normality (Complete Samples). **Biometrika**, Vol. 52, No. 3/4, p. 591 - 611, 1965.
- SPIEGEL, M. R. **Estatística**. São Paulo-Rio de Janeiro: Editora Mc Graw-Hill do Brasil, 1972. 580 p.
- STEEL, R. G. D.; TORRIE, J. H. **Principles and procedures of statistics: with special reference to the biological sciences**. New York: Mc Graw-Hill Book Company, Inc., 1960. 481 p.

CAPÍTULO 2

DELINEAMENTO INTEIRAMENTE AO ACASO

Este capítulo apresenta definições, conceitos e aplicações da análise de variância para experimentos com apenas um critério de classificação, isto é, o Delineamento Inteiramente Casualizado. Aborda ainda: as pressuposições básicas para a análise de variância; o desdobramento da soma de quadrados total e o teste F; os componentes de variância do modelo inteiramente ao acaso; as transformações de dados quando as variâncias forem heterogêneas; a análise de um ensaio inteiramente ao caso com K amostras por parcela; e a análise de covariância para o ensaio inteiramente ao acaso.

2.1. INTRODUÇÃO

Este delineamento experimental caracteriza-se por não apresentar nenhuma restrição quanto ao critério de aleatorização. As unidades experimentais que receberão os tratamentos são escolhidas completamente ao acaso, isto é, cada unidade experimental terá igual probabilidade de receber qualquer um dos tratamentos. Esse plano utiliza apenas os princípios da casualização e da repetição. Por essas características, ele é recomendável para ensaios em laboratórios e casas de vegetação, em que é possível exercer um eficiente controle das condições sobre as quais o experimento será realizado.

É um delineamento de fácil condução e de análise estatística mais simples, em que o número de tratamentos e de repetições depende tão somente do número de unidades experimentais disponíveis. Permite, inclusive, números diferentes de repetições entre os tratamentos. Logicamente, quanto maior for o número de graus de liberdade do resíduo, logo mais precisas serão as estimativas.

Para facilitar o entendimento, primeiramente o delineamento inteiramente ao acaso será apresentado com dados balanceados (mesmo número de repetições por tratamento) e, na sequência, com dados não balanceados (ausência da condição de mesmo número de repetições por tratamento).

O delineamento inteiramente ao acaso consiste em distribuir os tratamentos às unidades experimentais, por meio de um simples sorteio, isto é, sem qualquer tipo de restrição (equiprobabilidade). Isso significa, na teoria da probabilidade, que todos os resultados possíveis são igualmente prováveis, em que cada unidade experimental pode receber qualquer um dos tratamentos. Por outro lado, o procedimento de alocação dos tratamentos fica condicionado à disponibilidade de parcelas homogêneas no experimento.

O uso do delineamento inteiramente ao acaso é mais eficiente quanto maior for o grau de homogeneidade entre as unidades experimentais. Assim, se as unidades experimentais apresentarem heterogeneidade, o número de parcelas requeridas pode ser muito grande para obter a precisão desejada. Na ocorrência de heterogeneidade entre as parcelas experimentais, o ideal é estabelecer outros planos experimentais, como o delineamento casualizado em blocos ou o uso de variáveis auxiliares (covariáveis), objetivando diminuir o erro experimental.

O número de repetições é determinado a partir de informações da variabilidade das unidades experimentais relativo à variável dependente, do custo do ensaio e do poder dos testes de significância.

2.2. PRESSUPOSIÇÕES BÁSICAS PARA ANÁLISE DE VARIÂNCIA

A análise de variância necessita que algumas condições pressupostas sejam atendidas, pois caso não sejam obedecidas, então a validade das inferências realizadas é no mínimo questionável. As seguintes pressuposições básicas devem ser consideradas para realizar a análise de variância:

- a) Aditividade dos efeitos do modelo linear;
- b) Os erros são variáveis aleatórias independentes;
- c) Homogeneidade de variâncias dos tratamentos (variância constante);
- d) Distribuição normal ou aproximadamente normal dos erros.

Ferreira (2007) menciona que essas pressuposições, muitas vezes, não são averiguadas, o que pode comprometer a validade dos resultados dos testes, eventualmente podendo levar o pesquisador a tomar decisões errôneas. A maioria dos softwares disponíveis não verifica essas pressuposições, ou não possui rotinas para realização desses testes. Por outro lado, os testes existentes na literatura restringem-se a modelos específicos, o que dificulta a sua aplicação em casos mais gerais.

Ferreira (2007) também comenta que a pressuposição de normalidade, exigida na análise de variância, na maioria das vezes, é mal interpretada e verificada de forma incorreta. A exigência que se faz, a respeito da distribuição normal, é para a distribuição dos resíduos de um determinado modelo linear e não para os dados observados nas unidades experimentais. Muitos pesquisadores desavisados (ou por desconhecimento) realizam o teste de normalidade nos dados experimentais observados, o que é uma prática incorreta. Esse procedimento só seria válido se estivessem avaliando uma amostra aleatória de uma única população, cujos dados pudessem ser explicados pelo modelo linear simples $x_i = m + e_i$, em que x_i é a variável resposta, m é a média e e_i o erro.

Por outro lado, em modelos em que existe um ou mais fatores, os valores da variável resposta x são explicados por diferentes constantes ao longo da amostra aleatória de tamanho n . No modelo de um experimento inteiramente ao acaso $x_{ij} = m + t_i + e_{ij}$, em que ocorrem diferentes constantes ($m + t_i$) funções do i -ésimo nível do efeito t_i , a distribuição da variável resposta (x) é na verdade uma mistura de normais com diferentes médias. Assim sendo, quanto maior a complexidade do modelo, mais complicada ficará a mistura de distribuições normais.

Desse modo, como a suposição de normalidade é efetuada para o erro do modelo que é uma variável aleatória não observável, o mesmo tem que ser estimado para aplicação dos testes de normalidade.

Uma vez que a suposição de que os erros e_{ij} de um modelo têm distribuição normal e variância comum σ^2 , é possível indicar que as variâncias dos diferentes níveis dos fatores presentes no modelo devem ser homogêneas. No anexo B, é apresentado o programa **Prog6** (SAS), o qual testa a homogeneidade de variâncias do modelo inteiramente casualizado com um fator.

Em um modelo linear, é assumido que os efeitos são aditivos e não multiplicativos. O método de Tukey (1949a) decompõe a soma de quadrados dos erros em duas partes. Uma delas com apenas 1 grau de liberdade e a outra com os graus de liberdade restantes. Um teste F é aplicado e denominado de teste da não aditividade de Tukey. Esse teste pode ser generalizado para possibilitar sua aplicação em diversos modelos lineares. Tal generalização consiste em obter os valores preditos e, em seguida, introduzir o seu quadrado como covariável no modelo de análise de variância. Essa análise é aplicada unicamente para testar a hipótese de aditividade dos efeitos. Se houver efeito significativo da covariável, deve-se rejeitar a hipótese nula de efeitos aditivos.

Infelizmente esse procedimento não pode ser utilizado em experimentos inteiramente casualizados com um fator — por razões óbvias —, ou com dois fatores e interação, pois haverá, de acordo com Ferreira (2007), um confundimento da interação com o efeito da covariável.

No Capítulo 4 (exercício 4.2.1), está apresentada uma análise de um ensaio de competição de quatro espécies em cinco blocos, por meio do programa **Prog7**, anexado ao apêndice B, para a realização do teste de não aditividade de Tukey. Os programas **Prog5a** e **Prog5b** também são apresentados para efetuar os testes de normalidade de Shapiro-Wilk, Kolmogorov-Smirnov, Cramér-Von Mises e Anderson-Darling.

É importante efetuar a análise de resíduos, que consiste em construir um gráfico dos resíduos padronizados versus a ordem de coleta dos dados. Para a aceitação da pressuposição de independência, os resíduos devem ficar distribuídos casualmente em torno de zero sem um padrão definido.

$e_{ij} = x_{ij} - \hat{m}_i$: é o resíduo referente da j -ésima observação do i -ésimo tratamento,

$z_{ij} = \frac{e_{ij}}{\sqrt{QM_{RES}}}$: é o resíduo padronizado da j -ésima observação do i -ésimo tratamento, em que

QM_{RES} é o quadrado médio do resíduo.

2.3. DESDOBRAMENTO DA SOMA DE QUADRADOS TOTAL E O TESTE F

Tabela 2.3.1. Banco de dados com um fator fixo.

Trat.	Repetições				Totais
	1	2	...	J	
1	x_{11}	x_{12}	...	x_{1J}	$\sum_{j=1}^J x_{1j} = T_1$
2	x_{21}	x_{22}	...	x_{2J}	$\sum_{j=1}^J x_{2j} = T_2$
...	...	...	...	...	...
I	x_{I1}	x_{I2}	...	x_{IJ}	$\sum_{j=1}^J x_{ij} = T_I$
Total					$\sum_{i=1}^I \sum_{j=1}^J x_{ij} = G$

Seja considerado o modelo linear com apenas um critério de classificação:

$$x_{ij} = m + t_i + e_{ij} \quad e_{ij} = x_{ij} - m - t_i$$

Obtenção, pelo método dos quadrados mínimos, das estimativas dos parâmetros m e dos efeitos dos tratamentos t_i ($i = 1, 2, \dots, I$):

$$\sum_{i=1}^I \sum_{j=1}^J e_{ij}^2 = \sum_{i=1}^I \sum_{j=1}^J (x_{ij} - m - t_i)^2$$

$$z(m, t_i) = \sum_{i=1}^I \sum_{j=1}^J e_{ij}^2 = \sum_{i=1}^I \sum_{j=1}^J (x_{ij} - m - t_i)^2$$

Derivando em relação aos parâmetros m, t_i ($i = 1, 2, \dots, I$) e igualando a zero para obter o sistema de equações normais, tem-se:

$$\frac{\partial z(m, t_i)}{\partial m} = -2 \sum_{i=1}^I \sum_{j=1}^J (x_{ij} - \hat{m} - \hat{t}_i) = 0$$

$$\frac{\partial z(m, t_i)}{\partial t_i} = -2 \sum_{j=1}^J (x_{ij} - \hat{m} - \hat{t}_i) = 0$$

Resulta no sistema inconsistente seguinte que possui $I+1$ equações para $I+1$ incógnitas $(\hat{m}, \hat{t}_1, \hat{t}_2, \dots, \hat{t}_I)$. A soma das I equações de tratamentos é igual à equação da média, o que denota que as equações não são independentes e o sistema apresenta infinitas soluções.

$$\begin{cases} IJ\hat{m} + J \sum_{i=1}^I \hat{t}_i = \sum_{i=1}^I \sum_{j=1}^J x_{ij} = G & (1) \\ J\hat{m} + J\hat{t}_i = \sum_{j=1}^J x_{ij} = T_i & (2) \end{cases}$$

A solução de interesse é a que possibilita obter a estimativa da média independentemente do efeito de tratamento, isto é, estabelecendo a restrição:

$$\sum_{i=1}^I \hat{t}_i = 0 \quad \text{ou} \quad \sum_{i=1}^I J_i \hat{t}_i = 0 \quad (\text{ensaio desbalanceado})$$

Da equação (1), atendida a restrição imposta tem-se a estimativa da média ($\hat{m}$):

$$IJ\hat{m} = G \quad \rightarrow \quad \hat{m} = \frac{G}{IJ}$$

Da equação (2), tem-se a estimativa do efeito do i -ésimo tratamento ($\hat{t}_i$):

$$T_i - J\hat{m} = J\hat{t}_i \quad \hat{t}_i = \frac{T_i}{J} - \hat{m} = \hat{m}_i - \hat{m}$$

Por outro lado, obtidas as estimativas dos parâmetros do modelo matemático estabelecido, o próximo passo deverá ser a obtenção das somas de quadrados.

A soma de quadrados total corresponde à soma dos quadrados dos desvios de todas as observações em relação à média:

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J (x_{ij} - \hat{m})^2 = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}$$

$$C = \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = \frac{G^2}{IJ}$$

Considerando todas as IJ parcelas, a soma de quadrados de tratamentos ($SQ_{(TRAT)}$) corresponde à soma dos quadrados dos efeitos de todos os tratamentos.

$$SQ_{TRAT} = \hat{t}_1^2 + \hat{t}_1^2 + \dots + \hat{t}_1^2 + \hat{t}_2^2 + \hat{t}_2^2 + \dots + \hat{t}_2^2 + \dots + \hat{t}_I^2 + \hat{t}_I^2 + \dots + \hat{t}_I^2$$

$$SQ_{TRAT} = J\hat{t}_1^2 + J\hat{t}_2^2 + \dots + J\hat{t}_I^2 = J(\hat{t}_1^2 + \hat{t}_2^2 + \dots + \hat{t}_I^2)$$

Substituindo as estimativas dos efeitos $\hat{t}_i = \frac{T_i}{J} - \hat{m}$, em SQ_{TRAT} , tem-se:

$$SQ_{TRAT} = J[(\frac{T_1}{J} - \hat{m})^2 + (\frac{T_2}{J} - \hat{m})^2 + \dots + (\frac{T_I}{J} - \hat{m})^2]$$

$$SQ_{TRAT} = J(\frac{T_1^2}{J^2} - 2\hat{m}\frac{T_1}{J} + \hat{m}^2 + \frac{T_2^2}{J^2} - 2\hat{m}\frac{T_2}{J} + \hat{m}^2 + \dots + \frac{T_I^2}{J^2} - 2\hat{m}\frac{T_I}{J} + \hat{m}^2)$$

$$SQ_{TRAT} = \frac{T_1^2}{J} - 2\hat{m}T_1 + J\hat{m}^2 + \frac{T_2^2}{J} - 2\hat{m}T_2 + J\hat{m}^2 + \dots + \frac{T_I^2}{J} - 2\hat{m}T_I + J\hat{m}^2$$

$$SQ_{TRAT} = \frac{T_1^2}{J} + \frac{T_2^2}{J} + \dots + \frac{T_I^2}{J} - 2\hat{m}(T_1 + T_2 + \dots + T_I) + IJ\hat{m}^2$$

$$SQ_{TRAT} = \frac{1}{J}(T_1^2 + T_2^2 + \dots + T_I^2) - 2\frac{G}{IJ}G + IJ\frac{G^2}{I^2J^2}$$

$$SQ_{TRAT} = \frac{1}{J}(T_1^2 + T_2^2 + \dots + T_I^2) - \frac{G^2}{IJ}$$

$$SQ_{TRAT} = \frac{1}{J}(T_1^2 + T_2^2 + \dots + T_I^2) - C$$

$$SQ_{TRAT} = \frac{1}{J} \sum_{i=1}^I T_i^2 - C$$

A soma de quadrados do resíduo ($SQ_{RES.}$) corresponde ao efeito dos fatores não controlados.

$$SQ_{RES} = \sum_{I=1}^I \sum_{j=1}^J x_{ij}^2 - SQ_{PAR}(\hat{m}, \hat{t}_1, \hat{t}_2, \dots, \hat{t}_I)$$

A soma de quadrados de parâmetros ($SQ_{PAR.}$) é obtida pela soma dos produtos de cada parâmetro pelo respectivo segundo membro da equação normal, isto é:

De:

$$\begin{cases} IJ\hat{m} + J\sum_{i=1}^I \hat{t}_i = \sum_{i=1}^I \sum_{j=1}^J x_{ij} = G \\ J\hat{m} + \hat{t}_i = \sum_{j=1}^J x_{ij} = T_i \end{cases}$$

Tem-se:

$$SQ_{PAR} = \hat{m}G + \hat{t}_1 T_1 + \hat{t}_2 T_2 + \dots + \hat{t}_I T_I$$

$$SQ_{PAR} = \frac{G}{IJ}G + \left(\frac{T_1}{J} - \hat{m}\right)T_1 + \left(\frac{T_2}{J} - \hat{m}\right)T_2 + \dots + \left(\frac{T_I}{J} - \hat{m}\right)T_I$$

$$SQ_{PAR} = \frac{G^2}{IJ} + \frac{T_1^2}{J} - \hat{m}T_1 + \frac{T_2^2}{J} - \hat{m}T_2 + \dots + \frac{T_I^2}{J} - \hat{m}T_I$$

$$SQ_{PAR} = \frac{G^2}{IJ} + \frac{1}{J}(T_1^2 + T_2^2 + \dots + T_I^2) - \hat{m}(T_1 + T_2 + \dots + T_I)$$

$$SQ_{PAR} = \frac{G^2}{IJ} + \frac{1}{J}\sum_{i=1}^I T_i^2 - \frac{G}{IJ}G$$

$$SQ_{PAR} = \frac{G^2}{IJ} + \frac{1}{J}\sum_{i=1}^I T_i^2 - \frac{G^2}{IJ}$$

$$SQ_{PAR} = C + \frac{1}{J}\sum_{i=1}^I T_i^2 - C$$

Como:

$$SQ_{RES} = \sum_{I=1}^I \sum_{j=1}^J x_{ij}^2 - SQ_{PAR}(\hat{m}, \hat{t}_1, \hat{t}_2, \dots, \hat{t}_I)$$

Então:

$$SQ_{RES} = \sum_{I=1}^I \sum_{j=1}^J x_{ij}^2 - C - \frac{1}{J}\sum_{i=1}^I T_i^2 + C$$

$$SQ_{RES} = \left(\sum_{I=1}^I \sum_{j=1}^J x_{ij}^2 - C\right) - \left(\frac{1}{J}\sum_{i=1}^I T_i^2 - C\right)$$

Resultando em:

$$SQ_{RES} = SQ_{TOTAL} - SQ_{TRAT}$$

Considerando-se a homogeneidade das variâncias, ter-se-á:

$$QM_{RES} = \frac{\sum_{i=1}^I \sum_{j=1}^J (x_{ij} - \hat{m}_i)^2}{I(J-1)} = \frac{1}{I} \sum_{i=1}^I \left[\frac{\sum_{j=1}^J (x_{ij} - \hat{m}_i)^2}{J-1} \right] = \frac{1}{I} \sum_{i=1}^I \hat{\sigma}_i^2 = \hat{\sigma}^2$$

$$SQ_{TRAT} = \sum_{i=1}^I \sum_{j=1}^J (\hat{m}_i - \hat{m})^2 = J \sum_{i=1}^I (\hat{m}_i - \hat{m})^2 = \frac{\sum_{i=1}^I (\sum_{j=1}^J x_{ij})^2}{J} - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}$$

$$QM_{TRAT} = \frac{SQ_{TRAT}}{I-1}$$

Hipóteses a serem testadas:

$$H_0 : t_1 = t_2 = \dots = t_I \text{ vs } H_1 : t_i \neq t_u \text{ (} i \neq u \text{)}$$

$$H_0 : m_1 = m_2 = \dots = m_I \text{ (ou em que pelo menos duas médias são diferentes)}$$

H_0 : é denominada de hipótese de nulidade. Ela determina a ausência de efeito de tratamento, ou seja, a diferença entre tratamentos é não significante;

H_1 : definida como hipótese alternativa. Hipótese que indica se existe efeito de tratamentos ou se a diferença entre efeito de tratamentos é significante.

Quando se testa uma hipótese, em um determinado nível de significância α , rejeita-se ou não a hipótese de nulidade H_0 , podendo cometer um erro, isto é, rejeitar H_0 quando ela é verdadeira ou não rejeitar H_0 quando ela é falsa, o que pode ser resumido como segue:

Situação	Decisão	
	Aceitar H_0	Rejeitar H_0
H_0 : é verdadeira	Decisão correta	Erro tipo I
H_0 : é falsa	Erro tipo II	Decisão correta

A probabilidade de ocorrer um erro do tipo I ou erro de primeira espécie (Isto é, de aceitar como diferentes médias de tratamentos que na verdade são iguais), especificado pela letra grega α , e denominado de nível de significância. O erro do tipo I ou nível de significância, em um teste de hipótese, é normalmente fixado em 5% ou 1%. O erro tipo II ou erro de segunda espécie é cometido ao não rejeitar a hipótese H_0 quando ela é falsa, quer dizer, ao aceitar como iguais médias de tratamentos que são diferentes.

Estatística do teste:

$$F = \frac{QM_{TRAT}}{QM_{RES}}$$

Pelo teorema de Cochran (1934), tem-se:

- a) $\frac{SQ_{TRAT}}{\sigma^2}$: tem distribuição χ^2 com $(I - 1)$ graus de liberdade sob $H_0 : m_1 = m_2 = \dots = m_I$;
- b) $\frac{SQ_{RES}}{\sigma^2}$: possui distribuição χ^2 com $(n - I)$ graus de liberdade, tal que n é o número total de parcelas;
- c) SQ_{TRAT} e SQ_{RES} são independentes.

Logo, sob H_0 , a estatística F tem distribuição “ F de Snedecor” com $I - 1$ e $n - I$ graus de liberdade.

Rejeita-se H_0 , ao nível de significância α , se $F > F_{(I-1); \alpha}$

Exercício 2.3.1. Seja considerado um ensaio inteiramente ao acaso referente à produção de madeira de uma espécie florestal, a qual foi submetida a quatro diferentes tipos de manejo: M1, M2, M3 e M4.

Tabela 2.3.2. Volume de madeira em m³ por parcela (Dados simulados).

Repetições	M1	M2	M3	M4
1	34,5	45,0	43,5	42,0
2	38,3	53,3	54,0	36,8
3	43,5	54,0	51,8	44,3
4	41,3	51,0	50,3	42,8
5	39,5	49,6	53,4	43,0
6	44,0	52,3	45,6	34,1
Totais	241,1	305,2	298,6	243,0
Médias	40,1833	50,8667	49,7667	40,5000

$$SQ_{TOTAL} = \sum_{I=1}^I \sum_{j=1}^J x_{ij}^2 - C = 34,5^2 + 38,3^2 + \dots + 43,0^2 + 34,1^2 - C = 50.206,55 - C$$

$$C = \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = \frac{(34,5 + 38,3 + \dots + 43,0 + 34,1)^2}{24} = \frac{1.087,9^2}{24} = 49.313,6004$$

$$SQ_{TOTAL} = \sum_{I=1}^I \sum_{j=1}^J x_{ij}^2 - C = 50.206,55 - 49.313,6004 = 892,9496$$

$$SQ_{TRAT} = \frac{1}{J} \sum_{i=1}^I T_i^2 - C = \frac{1}{6} (241,1^2 + 305,2^2 + 298,6^2 + 243,0^2) - 49.313,6004$$

$$SQ_{TRAT} = \frac{1}{6} \times 299.487,21 - 49.313,6004 = 600,9346$$

$$SQ_{RES} = SQ_{TOTAL} - SQ_{TRAT} = 892,9496 - 600,9346 = 292,0150$$

Resultando na seguinte tabela de análise de variância (ANOVA):

Tabela 2.3.3. ANOVA.

F.V. (Fontes de Variação)	G.L.	S.Q.	Q.M.	<i>F</i>
Tratamentos	3	600,9346	200,3115	13,72**
Resíduo	20	292,0150	14,6007	
Total	23	892,9496		

$$F_{(I-1);(n-I);\alpha} = F_{3;20;0,05} = 3,10$$

Como $F > F_{3;20;0,05}$, então se rejeita $H_0 : m_1 = m_2 = m_3 = m_4$ ao nível $\alpha = 0,05$ de significância.

Pode ocorrer, em decorrência do planejamento do experimento, que o número de unidades experimentais disponível não seja múltiplo do número de tratamentos. Ou pode acontecer que o experimento tenha sido implantado com os dados balanceados, mas que, devido a problemas ambientais, algumas parcelas experimentais tenham sido perdidas. Então, conseqüentemente, os tratamentos podem ficar com números de repetições total ou parcialmente diferentes, ou seja, o experimento se tornou não balanceado (número diferente de repetições).

O ideal é que todos os tratamentos apresentem o mesmo número de repetições, possibilitando analisar o experimento pelo método exato. Entretanto, não sendo possível, deve ser adotada a metodologia de análise de dados balanceados, obviamente, adaptando as fórmulas dos cálculos aos experimentos com dados não balanceados.

Esse tipo de análise direciona o processo exato para uma forma aproximada. Por outro lado, a teoria desenvolvida para o ensaio inteiramente balanceado passa a ser explicitada para os dados não balanceados, com pequenas modificações nas somas de quadrados e nos procedimentos de comparações múltiplas, como, por exemplo, no teste de Tukey.

Exercício 2.3.2. Seja considerado o ensaio inteiramente ao acaso não balanceado referente à produção de madeira de uma espécie florestal, a qual foi submetida a quatro diferentes tipos de manejo: M1, M2, M3 e M4.

Tabela 2.3.4. Volume de madeira (m³ por parcela)

Repetições	M1	M2	M3	M4
1	34,5	45,0	43,5	42,0
2	38,3	53,3	54,0	36,8
3	43,5	54,0	51,8	44,3
4	41,3	51,0	50,3	42,8
5	39,5	49,6	---	43,0
6	44,0	52,3	---	---
Totais	241,1	305,2	199,6	208,9
Médias	40,18	50,87	49,90	41,78

$$\sum_{I=1}^I \sum_{j=1}^{J_i} x_{ij}^2 = 34,5^2 + 38,3^2 + \dots + 42,8^2 + 43,0^2 = 44.112,82$$

$$C = \frac{(\sum_{i=1}^I T_i)^2}{21} = \frac{G^2}{21} = \frac{(241,1 + 305,2 + 199,6 + 208,9)^2}{21} = \frac{954,8^2}{21} = 43.411,5733$$

$$SQ_{TOTAL} = \sum_{I=1}^I \sum_{j=1}^{J_i} x_{ij}^2 - C = 44.112,82 - 43.411,5733 = 701,2467$$

$$SQ_{TRAT} = \left(\frac{T_1^2}{J_1} + \frac{T_2^2}{J_2} + \frac{T_3^2}{J_3} + \frac{T_4^2}{J_4} \right) - C = \frac{241,1^2}{6} + \frac{305,2^2}{6} + \frac{199,6^2}{4} + \frac{208,9^2}{5} - C$$

$$SQ_{TRAT} = \left(\frac{T_1^2}{J_1} + \frac{T_2^2}{J_2} + \frac{T_3^2}{J_3} + \frac{T_4^2}{J_4} \right) - C = 43.900,5903 - 43.411,5733 = 489,0170$$

$$SQ_{RES} = SQ_{TOTAL} - SQ_{TRAT} = 701,25 - 489,02 = 212,23$$

Tabela 2.3.5. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	3	489,0170	163,0057	13,06**
Resíduo	17	212,2297	12,4841	
Total	20	701,2467		

$$F_{(I-1); (n-I); \alpha} = F_{3; 17; 0,05} = 3,20 \quad F_{3; 17; 0,01} = 5,18$$

Considerando $H_0 = m_1 = m_2 = m_3 = m_4$ e como $F > F_{3; 20; 0,05}$, portanto se rejeita H_0 ao nível de significância $\alpha = 0,05$.

2.4. COMPONENTES DE VARIÂNCIA DO MODELO INTEIRAMENTE AO ACASO.

Os componentes de variância são definidos como as variâncias associadas aos efeitos aleatórios de um modelo matemático. De acordo com Barbin (2019), existem os seguintes métodos para obter os componentes de variância:

- i) Método dos momentos ou da análise de variância;
 - ii) Método da máxima verossimilhança; os métodos de Henderson I, II e III de análise de variância para dados não balanceados;
 - iii) Método da máxima verossimilhança restrita;
 - iv) Métodos de Rao (MINQUE-Minimum Norm Quadratic Unbiased Estimative e MIVQUE-Minimum Quadratic Unbiased Estimative);
 - v) Métodos de Bayes.
- a) Segue o modelo matemático para um ensaio inteiramente ao acaso, a fim de obter os componentes de variância, considerando o modelo aleatório, pelo método dos momentos ou de análise de variância:

$$x_{ij} = m + t_i + e_{ij} \quad \begin{cases} i = 1, 2, \dots, I \\ j = 1, 2, \dots, J \end{cases}$$

m : média geral (fixa), tal que $E(m) = m$ e $E(m^2) = m^2$;

t_i : efeito aleatório do i -ésimo tratamento, em que $t_i \sim NID(0, \sigma_t^2)$, ou seja, normal e independentemente distribuído, com média zero e variância σ_t^2 , então $E(t_i) = 0$ e $E(t_i^2) = \sigma_t^2$.

e_{ij} : erro experimental, considerado aleatório, normal e independentemente distribuído, ou seja, $e_{ij} \sim NID(0, \sigma^2)$, então $E(e_{ij}) = 0$ e $E(e_{ij}^2) = \sigma^2$.

Considera-se, para obtenção dos componentes de variância do modelo aleatório, que os efeitos aleatórios do modelo não sejam correlacionados, logo:

$$Cov(t_i, e_{ij}) = 0 \text{ para } (\forall i, j);$$

$$Cov(t_i, t_{i'}) = 0 \text{ para } (i \neq i')$$

$$Cov(e_{ij}, e_{i'j'}) = 0 \text{ para } (i \neq i' \text{ e } j \neq j')$$

Para a soma dos quadrados total (SQ_{TOTAL}):

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}$$

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C \quad C = \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}$$

$$E[\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2] = \sum_{i=1}^I \sum_{j=1}^J E[(m + t_i + e_{ij})^2]$$

$$E[\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2] = \sum_{i=1}^I \sum_{j=1}^J E[m^2 + t_i^2 + e_{ij}^2 + 2mt_i + 2me_{ij} + 2t_ie_{ij}]$$

$$E[\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2] = \sum_{i=1}^I \sum_{j=1}^J (m^2 + \sigma_t^2 + \sigma^2) = IJm^2 + IJ\sigma_t^2 + IJ\sigma^2$$

$$E[C] = E[\frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}] = \frac{1}{IJ} E\{[\sum_{i=1}^I \sum_{j=1}^J (m + t_i + e_{ij})]^2\}$$

$$E[C] = \frac{1}{IJ} E[I^2 J^2 m^2 + J^2 (\sum_{i=1}^I t_i)^2 + (\sum_{i=1}^I \sum_{j=1}^J e_{ij})^2 + dp]$$

dp : duplos produtos

Em que:

$$E[(\sum_{i=1}^I t_i)^2] = E[(t_1 + t_2 + \dots + t_I)^2]$$

$$E[(\sum_{i=1}^I t_i)^2] = E[t_1^2 + t_2^2 + \dots + t_I^2 + 2t_1t_2 + \dots + 2t_{I-1}t_I]$$

$$E[(\sum_{i=1}^I t_i)^2] = \sigma_t^2 + \sigma_t^2 + \dots + \sigma_t^2 + 0 + \dots + 0 = I\sigma_t^2$$

Assim como:

$$E[(\sum_{i=1}^I \sum_{j=1}^J e_{ij})^2] = E[(e_{11} + e_{12} + \dots + e_{IJ})^2]$$

$$E[(\sum_{i=1}^I \sum_{j=1}^J e_{ij})^2] = E[e_{11}^2 + e_{12}^2 + \dots + e_{1J}^2 + dp + e_{21}^2 + e_{22}^2 + \dots + e_{2J}^2 + dp + \dots + \dots + \dots + \dots + e_{I1}^2 + e_{I2}^2 + \dots + e_{IJ}^2 + dp] = IJ\sigma^2$$

Logo:

$$E[C] = \frac{1}{IJ} E[I^2 J^2 m^2 + J^2 (\sum_{i=1}^I t_i)^2 + (\sum_{i=1}^I \sum_{j=1}^J e_{ij})^2 + dp \text{ (duplos produtos)}]$$

$$E[C] = \frac{1}{IJ} [I^2 J^2 m^2 + J^2 I\sigma_t^2 + IJ\sigma^2]$$

$$E[C] = IJm^2 + J\sigma_t^2 + \sigma^2$$

Bem como:

$$E[\sum_{I=1}^I \sum_{J=1}^J x_{ij}^2] = \sum_{I=1}^I \sum_{J=1}^J [m^2 + \sigma_t^2 + \sigma^2] = IJm^2 + IJ\sigma_t^2 + IJ\sigma^2$$

Logo:

$$E[SQ_{TOTAL}] = IJm^2 + IJ\sigma_t^2 + IJ\sigma^2 - (IJm^2 + J\sigma_t^2 + \sigma^2)$$

$$E[SQ_{TOTAL}] = IJ\sigma_t^2 + IJ\sigma^2 - J\sigma_t^2 - \sigma^2$$

$$E[SQ_{TOTAL}] = \sigma_t^2(IJ - J) + \sigma^2(IJ - 1)$$

$$E(QM_{TOTAL}) = \frac{\sigma_t^2(IJ - J) + \sigma^2(IJ - 1)}{(IJ - 1)}$$

$$E(QM_{TOTAL}) = \frac{J(I - 1)\sigma_t^2 + (IJ - 1)\sigma^2}{(IJ - 1)}$$

Para a soma de quadrados dos tratamentos (SQ_{TRAT}):

$$SQ_{TRAT} = \frac{1}{J} \sum_{i=1}^I T_i^2 - C = \frac{1}{J} \sum_{i=1}^I x_{i.}^2 - C \quad x_{i.} = \sum_{j=1}^J x_{ij}$$

$$E\left[\frac{1}{J} \sum_{i=1}^I x_{i.}^2\right] = \frac{1}{J} E\left[\sum_{i=1}^I x_{i.}^2\right] = \frac{1}{J} \sum_{i=1}^I E(x_{i.}^2)$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = \frac{1}{J} \sum_{i=1}^I E\left[(Jm + Jt_i + \sum_{j=1}^J e_{ij})^2\right]$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = \frac{1}{J} \sum_{i=1}^I E[J^2 m^2 + J^2 t_i^2 + (\sum_{j=1}^J e_{ij})^2 + dp]$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = \frac{1}{J} (IJ^2 m^2 + IJ^2 \sigma_t^2 + IJ\sigma^2 + 0)$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = IJm^2 + IJ\sigma_t^2 + I\sigma^2$$

Dado que:

$$E[C] = IJm^2 + J\sigma_t^2 + \sigma^2$$

Então:

$$E[SQ_{TRAT}] = IJm^2 + IJ\sigma_t^2 + I\sigma^2 - (IJm^2 + J\sigma_t^2 + \sigma^2)$$

$$E[SQ_{TRAT}] = IJ\sigma_t^2 + I\sigma^2 - J\sigma_t^2 - \sigma^2$$

$$\boxed{E[SQ_{TRAT}] = (I-1)\sigma^2 + J(I-1)\sigma_t^2}$$

$$E(QM_{TRAT}) = \frac{(I-1)\sigma^2 + J(I-1)\sigma_t^2}{I-1} = \sigma^2 + J\sigma_t^2$$

Para a soma de quadrados dos resíduos (SQ_{RES}):

$$SQ_{RES} = SQ_{TOTAL} - SQ_{TRAT}$$

$$E(SQ_{RES}) = \sigma_t^2 (IJ - J) + \sigma^2 (IJ - 1) - [(I-1)\sigma^2 + J(I-1)\sigma_t^2]$$

$$\boxed{E(SQ_{RES}) = I(J-1)\sigma^2}$$

$$E(QM_{RES}) = \frac{I(J-1)\sigma^2}{I(J-1)} = \sigma^2$$

Resultando no quadro de análise de variância (tabela 2.4.1) para o delineamento inteiramente ao acaso (modelo aleatório).

Tabela 2.4.1. Componentes de variância (modelo aleatório).

F.V.	G.L.	E(S.Q.)	E(Q.M.)	F
Trat.	$I - 1$	$J(I - 1)\sigma_t^2 + (I - 1)\sigma^2$	$\sigma^2 + J\sigma_t^2$	$\frac{\sigma^2 + J\sigma_t^2}{\sigma^2}$
Resíduo	$I(J - 1)$	$I(J - 1)\sigma^2$	σ^2	
Total	$IJ - 1$	$J(I - 1)\sigma_t^2 + (IJ - 1)\sigma^2$	$\frac{J(I - 1)\sigma_t^2 + (IJ - 1)\sigma^2}{IJ - 1}$	

No caso do modelo aleatório, a hipótese a ser testada será: $H_0 : \sigma_t^2 = 0$, isto é, a variância dos efeitos dos tratamentos é nula.

b) Segue o modelo matemático para o ensaio inteiramente ao acaso (modelo fixo), em que a obtenção dos componentes de variância ocorre pelo método dos momentos ou de análise de variância:

$$x_{ij} = m + t_i + e_{ij} \quad \begin{cases} i = 1, 2, \dots, I \\ j = 1, 2, \dots, J \end{cases}$$

m : média geral (fixa), tal que $E(m) = m$ e $E(m^2) = m^2$;

t_i : efeito fixo do i -ésimo tratamento com $E(t_i) = t_i$, $E(t_i^2) = t_i^2$ e $\sum_{i=1}^I t_i = 0$.

e_{ij} : erro experimental, considerado aleatório, normal e independentemente distribuído, ou seja, $e_{ij} \sim NID(0, \sigma^2)$, então $E(e_{ij}) = 0$, $E(e_{ij}^2) = \sigma^2$ e $Cov(e_{ij}, e_{i'j'}) = 0$ para $(i \neq i' \text{ e } j \neq j')$.

Para a soma dos quadrados total (SQ_{TOTAL}):

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}$$

$$E[\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2] = \sum_{i=1}^I \sum_{j=1}^J E[(m + t_i + e_{ij})^2]$$

$$E[\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2] = \sum_{i=1}^I \sum_{j=1}^J E[m^2 + t_i^2 + e_{ij}^2 + 2mt_i + \text{outros } dp]$$

$$E[\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2] = \sum_{i=1}^I \sum_{j=1}^J (m^2 + t_i^2 + \sigma^2 + 2mt_i)$$

$$E[\sum_{I=1}^I \sum_{j=1}^J x_{ij}^2] = IJm^2 + J \sum_{i=1}^I t_i^2 + IJ\sigma^2 + 2Jm \sum_{i=1}^I t_i$$

$$E[\sum_{I=1}^I \sum_{j=1}^J x_{ij}^2] = IJm^2 + J \sum_{i=1}^I t_i^2 + IJ\sigma^2$$

$$E[C] = E[\frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}] = \frac{1}{IJ} E\{[\sum_{i=1}^I \sum_{j=1}^J (m + t_i + e_{ij})]^2\}$$

$$E[C] = E[\frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}] = \frac{1}{IJ} E[(IJm + J \sum_{i=1}^I t_i + \sum_{i=1}^I \sum_{j=1}^J e_{ij})^2] \quad \sum_{i=1}^I t_i = 0$$

$$E[C] = E[\frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}] = \frac{1}{IJ} E[I^2 J^2 m^2 + (\sum_{i=1}^I \sum_{j=1}^J e_{ij})^2 + 2IJm \sum_{i=1}^I \sum_{j=1}^J e_{ij}] \quad \sum_{i=1}^I \sum_{j=1}^J e_{ij} = 0$$

Dado que:

$$E[(\sum_{i=1}^I \sum_{j=1}^J e_{ij})^2] = IJ\sigma^2$$

Então:

$$E[C] = E[\frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}] = \frac{1}{IJ} (I^2 J^2 m^2 + IJ\sigma^2)$$

$$E[C] = IJm^2 + \sigma^2$$

Como:

$$E[\sum_{I=1}^I \sum_{j=1}^J x_{ij}^2] = IJm^2 + J \sum_{i=1}^I t_i^2 + IJ\sigma^2$$

Então:

$$E[SQ_{TOTAL}] = IJm^2 + J \sum_{i=1}^I t_i^2 + IJ\sigma^2 - (IJm^2 + \sigma^2)$$

Logo:

$$E[SQ_{TOTAL}] = J \sum_{i=1}^I t_i^2 + IJ\sigma^2 - \sigma^2$$

$$E[QM_{TOTAL}] = \frac{J \sum_{i=1}^I t_i^2 + IJ\sigma^2 - \sigma^2}{IJ - 1}$$

Soma de quadrados dos tratamentos (SQ_{TRAT}):

$$SQ_{TRAT} = \frac{1}{J} \sum_{i=1}^I T_i^2 - C = \frac{1}{J} \sum_{i=1}^I x_{i.}^2 - C \quad x_{i.} = \sum_{j=1}^J x_{ij}$$

$$E\left[\frac{1}{J} \sum_{i=1}^I x_{i.}^2\right] = \frac{1}{J} E\left[\sum_{i=1}^I x_{i.}^2\right] = \frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)]$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = \frac{1}{J} \sum_{i=1}^I \{E[(Jm + Jt_i + \sum_{j=1}^J e_{ij})^2]\}$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = \frac{1}{J} \sum_{i=1}^I \{E[J^2 m^2 + J^2 t_i^2 + (\sum_{j=1}^J e_{ij})^2 + 2J^2 m t_i + \text{outros } dp]\}$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = \frac{1}{J} \sum_{i=1}^I (J^2 m^2 + J^2 t_i^2 + J\sigma^2 + 2J^2 m t_i) \quad E(\sum_{j=1}^J e_{ij})^2 = J\sigma^2$$

$$\frac{1}{J} \sum_{i=1}^I [E(x_{i.}^2)] = IJm^2 + J \sum_{i=1}^I t_i^2 + I\sigma^2$$

Como:

$$E[C] = E\left[\frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}\right] = IJm^2 + \sigma^2$$

Então:

$$E[SQ_{TRAT}] = IJm^2 + J \sum_{i=1}^I t_i^2 + I\sigma^2 - (IJm^2 + \sigma^2)$$

$$E[SQ_{TRAT}] = J \sum_{i=1}^I t_i^2 + I\sigma^2 - \sigma^2 = J \sum_{i=1}^I t_i^2 + \sigma^2 (I-1)$$

$$E[SQ_{TRAT}] = J \sum_{i=1}^I t_i^2 + \sigma^2 (I-1)$$

$$E(QM_{TRAT}) = \frac{J \sum_{i=1}^I t_i^2 + \sigma^2 (I-1)}{I-1} = \frac{J \sum_{i=1}^I t_i^2}{I-1} + \sigma^2$$

$$E(SQ_{RES}) = SQ_{TOTAL} - SQ_{TRAT}$$

$$E(SQ_{RES}) = J \sum_{i=1}^I t_i^2 + IJ\sigma^2 - \sigma^2 - [J \sum_{i=1}^I t_i^2 + \sigma^2(I-1)]$$

$$E(SQ_{RES}) = IJ\sigma^2 - \sigma^2 - \sigma^2(I-1)$$

$$E(SQ_{RES}) = \sigma^2(IJ - I - I + 1) = I(J-1)\sigma^2$$

$$E(QM_{RES}) = \frac{I(J-1)\sigma^2}{I(J-1)} = \sigma^2$$

Tabela 2.4.2. Componentes de variância do modelo fixo.

F.V.	G.L.	E(S.Q.)	E(Q.M.)	F
Trat.	$I - 1$	$J \sum_{i=1}^I t_i^2 + (I-1)\sigma^2$	$\sigma^2 + \frac{J \sum_{i=1}^I t_i^2}{I-1}$	$\frac{\sigma^2 + J \sum_{i=1}^I t_i^2}{\sigma^2}$
Resíduo	$I(J-1)$	Diferença	σ^2	
Total	$IJ - 1$	$J \sum_{i=1}^I t_i^2 + (IJ-1)\sigma^2$	$\frac{J \sum_{i=1}^I t_i^2 + (IJ-1)\sigma^2}{IJ-1}$	

No caso do modelo fixo, a hipótese a ser testada, será:

$H_0 : t_1 = t_2 = \dots = t_I$, ou, $H_0 : m_1 = m_2 = \dots = m_I$, isto é, não significância entre os efeitos dos tratamentos.

c) Componentes de variância do ensaio inteiramente casualizado com K amostras por parcela (modelo aleatório).

$$x_{ijk} = m + t_i + e_{ij} + \varepsilon_{ijk}$$

m : média geral;

t_i : efeito aleatório de tratamento, normalmente distribuído com média zero e σ_t^2 , em que $E(t_i) = 0$ e $E(t_i^2) = \sigma_t^2$;

e_{ij} : erro experimental (erro entre parcelas dentro de tratamentos);

$\varepsilon_{(ij)k}$: erro amostral (entre indivíduos dentro das parcelas).

Tabela 2.4.3. Componentes de variância do ensaio inteiramente casualizado com K amostras por parcela (modelo aleatório).

F.V.	G.L.	E(Q.M.)
Tratamentos	$I - 1$	$\sigma_\varepsilon^2 + K\sigma_e^2 + JK\sigma_t^2$
Resíduo	$I(J-1)$	$\sigma_\varepsilon^2 + K\sigma_e^2$
(Parcelas)	$IJ - 1$	
Entre indivíduos d. Parcelas	$IJ(K-1)$	σ_ε^2
Total	$IJK-1$	

d) Componentes de variância do ensaio inteiramente casualizado com K amostras por parcelas (modelo fixo):

m : média geral;

t_i : efeito fixo com $E(t_i) = t_i$, $E(t_i^2) = t_i^2$ e $\sum_{i=1}^I t_i = 0$;

e_{ij} : erro experimental (erro entre parcelas dentro de tratamentos);

$\mathcal{E}_{(ij)k}$: erro amostral (entre indivíduos dentro das parcelas).

Tabela 2.4.4. Componentes de variância do ensaio inteiramente casualizado com K amostras por parcela (modelo fixo).

F.V.	G.L.	E(Q.M.)
Trat.	$I - 1$	$\sigma_{\varepsilon}^2 + K\sigma_e^2 + JK\phi_t$
Resíduo	$I(J-1)$	$\sigma_{\varepsilon}^2 + K\sigma_e^2$
(Parcelas)	$IJ - 1$	
Entre indivíduos d. Parcelas	$IJ(K-1)$	σ_{ε}^2
Total	$IJK-1$	

$$\phi_t = \frac{\sum_{i=1}^I t_i^2}{I}$$

e) Método prático de Hicks (1973) para o caso dos modelos balanceados, que consiste em estabelecer uma tabela de dupla entrada e, de acordo com Barbin (2019), seguir as seguintes etapas:

Seja considerado o ensaio inteiramente casualizado com K amostras por parcela (modelo aleatório):

$$x_{ijk} = m + t_i + e_{ij} + \mathcal{E}_{ijk}$$

e1) Colocar à esquerda as fontes especificadas pelos próprios efeitos do modelo.

t_i	
e_{ij}	
$\mathcal{E}_{(ij)k}$	

e2) Colocar na linha superior da tabela os índices referidos no modelo, em que I, J e K correspondem ao número de observações.

Logo:

	I a_i	J a_j	K a_k
t_i			
e_{ij}			
$\mathcal{E}_{(ij)k}$			

A letra (a) indica se tratar de um modelo aleatório, enquanto que a letra (f) corresponde ao modelo fixo.

$\mathcal{E}_{(ij)k}$: (ij) entre parênteses significa variação “dentro de”.

e3) Em cada coluna, inserir o número de observações (I , J e K) nas linhas em que o índice da coluna não aparece, ou seja, colocar I (maiúsculo) onde não há i (minúsculo), J onde não há j (minúsculo) e K onde não há k (minúsculo).

	I a_i	J a_j	K a_k
t_i		J	K
e_{ij}			K
$\mathcal{E}_{(ij)k}$			

e4) Nas linhas em que aparecem índices entre parênteses, inserir 1 nas colunas referentes a esses índices.

	I a_i	J a_j	K a_k
t_i		J	K
e_{ij}			K
$\mathcal{E}_{(ij)k}$	1	1	

e5) Colocar o número um (1), nas casas restantes, se o efeito é aleatório, isto é, se aparecer a letra (a), mas, se ocorrer a letra (f), colocar zero (0).

	I a_i	J a_j	K a_k
t_i	1	J	K
e_{ij}	1	1	K
$\mathcal{E}_{(ij)k}$	1	1	1

e6) Finalizando, para obter os $E(QM)$, é necessário:

1) cobrir as colunas correspondentes aos índices que ocorrem na linha cujo objetivo é obter a $E(QM)$. Por outro lado, não se cobrem as colunas quando existirem índices entre parênteses;

2) multiplicar em cada linha os valores internos da tabela. Esses produtos são multiplicados, respectivamente, pela variância do efeito. A soma desses produtos corresponderá a $E(QM)$. Entretanto, devem ser eliminadas dessa soma todas as parcelas oriundas de linhas que não contenham o índice do efeito a que se refere a $E(QM)$.

No caso da $E(QM_{TRAT})$ cobrindo a coluna (a_i^I), então, na linha (t_i), tem-se: $J \times K \times \sigma_t^2$, na linha e_{ij} : $1 \times K \times \sigma_e^2$ e na linha $\varepsilon_{(ij)k}$: $1 \times 1 \times \sigma_\varepsilon^2$. A soma dos produtos será: $\sigma_\varepsilon^2 + K\sigma_e^2 + JK\sigma_t^2$.

No caso da $E(QM_e)$ cobrindo as colunas (a_j^J) e (a_k^J), então, na linha (t_i), tem-se: $K \times \sigma_t^2$ (Eliminado, pois não contém o índice “j”), na linha e_{ij} : $K \times \sigma_e^2$, na linha $\varepsilon_{(ij)k}$: $1 \times \sigma_\varepsilon^2$. A soma dos produtos será: $\sigma_\varepsilon^2 + K\sigma_e^2$.

No caso da $E(QM_\varepsilon)$ não se cobrem as colunas, pois há índices entre parênteses. Na linha (t_i), tem-se: $1 \times J \times K \times \sigma_t^2$ (Eliminado, pois (t_i) não contém o índice “k”), na linha e_{ij} : $1 \times 1 \times K \times \sigma_e^2$ (Eliminado, pois (e_{ij}) não contém o índice “k”), na linha $\varepsilon_{(ij)k}$: $1 \times 1 \times 1 \times \sigma_\varepsilon^2$. Resultado da soma: σ_ε^2 .

Então, considerando o ensaio inteiramente casualizado com K amostras por parcela (modelo aleatório), os componentes de variâncias são os seguintes:

	a_i^I	a_j^J	a_k^K	$E(QM)$
t_i	1	J	K	$\sigma_\varepsilon^2 + K\sigma_e^2 + JK\sigma_t^2$
e_{ij}	1	1	K	$\sigma_\varepsilon^2 + K\sigma_e^2$
$\varepsilon_{(ij)k}$	1	1	1	σ_ε^2

2.5. TRANSFORMAÇÕES DE DADOS NO CASO DE HETEROGENEIDADE DE VARIÂNCIAS

A transformação de dados é um procedimento empregado quando as variâncias forem heterogêneas, sendo que o tipo de transformação deve ser escolhido segundo critérios definidos.

A seguir sejam as seguintes transformações: Raiz quadrada, Logarítmica e Arco seno ou angular:

a) Raiz quadrada

A transformação pela raiz quadrada é indicada para dados explicados pela distribuição de Poisson, de modo que a média é aproximadamente igual à variância e os dados apresentam probabilidade muito baixa de ocorrer. Recomenda-se para dados oriundos de contagem: número de sementes por parcela, número de insetos por planta, número de árvores danificadas por parcela, dados provenientes de uma escala de notas etc.

b) Logarítmica

Aplica-se a logaritmização dos dados quando os desvios padrões das amostras são aproximadamente proporcionais à média, ou seja, as amostras apresentam relativamente os mesmos coeficientes de variação. Essa transformação é recomendada também quando os efeitos dos tratamentos são multiplicativos. Na transformação logarítmica, as variâncias de cada amostra devem ser obtidas com no mínimo 15 observações.

c) Arco seno ou angular

É a transformação empregada para dados com a média proporcional à variância, isto é, dados provenientes de uma distribuição binomial, como, por exemplo, o caso de dados originados de proporção, tal como porcentagem de sementes germinadas.

A forma mais simples de escolher a melhor transformação é por meio de gráficos. Neles, colocam-se, nos eixos dos x e y , as médias e as variâncias de cada amostra, considerando cada transformação. Posteriormente seleciona-se a que apresentar a menor dispersão. Outra forma é calcular para o maior e o menor dado de cada amostra a sua amplitude correspondente, obtendo a razão entre a maior e a menor amplitude, escolhendo, nesse caso, a transformação que apresente a menor amplitude.

Uma indicação razoável é o Coeficiente de Variação (CV) dos dados transformados. Quanto menor o CV, melhor será em relação aos dados originais.

Os testes de comparações múltiplas são realizados na escala da variável transformada, mas se o pesquisador preferir apresentar os dados (médias) na escala original, obrigatoriamente deverá utilizar o inverso da transformação utilizada.

A heteroscedasticidade (Não homogeneidade de variância) regular é verificada pela falta de normalidade dos dados experimentais, denotando, frequentemente, certa relação entre a média e a variância dos diversos tratamentos.

A heteroscedasticidade irregular ocorre quando alguns tratamentos apresentam maior variabilidade que outros. Por exemplo, nos experimentos com inseticidas, nos quais é considerado um grupo de parcelas não tratadas (Testemunha ou controle). Nesses experimentos, é comum a situação em que o número de insetos nas parcelas tratadas é menor e mais homogêneo do que na testemunha, a qual apresenta maior variabilidade.

Constatada a não homogeneidade das variâncias e considerada a existência de uma heteroscedasticidade regular, procede-se à transformação logarítmica, uma vez que existe uma proporcionalidade entre a média e o desvio padrão dos tratamentos.

Para melhor definir o tipo de transformação a ser empregada, Box e Cox (1964) recomendam estabelecer uma regressão linear entre $\log \hat{\sigma}^2$ como variável dependente e $\log \hat{m}$ como independente, isto é, $\log \hat{\sigma}^2 = \hat{a} + \hat{b} \log \hat{m}$. O método consiste em determinar o valor de λ (tabela 2.5.1), que definirá qual o tipo de transformação a ser empregado. Tal que $\lambda = 1 - \frac{\hat{b}}{2}$.

Para $\lambda \neq 0$, então $Y^* = Y^\lambda$, mas se $\lambda = 0$, então $Y^* = \log Y$.

Tabela 2.5.1. Alternativas de transformação de dados.

$\hat{b}$	λ	Tipo de transformação
0	1	nenhuma
1	0,5	$\sqrt{X}$ ou $\sqrt{Y}$
2	0	$\log X$ ou $\log Y$
3	-0,5	$1/\sqrt{X}$ ou $1/\sqrt{Y}$
4	-1	$1/X$ ou $1/Y$

Exercício 2.5.1. Testar a homogeneidade de variâncias pelos testes da razão máxima de Hartley, Bartlett e Cochran, por meio de um experimento inteiramente casualizado com cinco tratamentos e seis repetições (dados simulados), visando verificar a eficiência no controle de uma determinada praga em um viveiro florestal.

Tabela 2.5.2. Número de indivíduos mortos após 48 horas.

Repetições	Tratamentos				
	A	B	C	D	E
1	5.546,0	1.648,0	926,0	208,0	158,0
2	5.406,0	1.692,0	892,0	198,0	189,0
3	5.233,0	1.845,0	857,0	268,0	207,0
4	5.646,0	1.810,0	801,0	189,0	142,0
5	5.412,0	1.536,0	667,0	167,0	172,0
6	5.623,0	1.541,0	773,0	134,0	168,0
Variâncias	24.687,5	17.067,9	8.753,1	2.008,4	524,7

a) Teste da razão máxima de Hartley

$$H = \frac{s_{Máx}^2}{s_{Min}^2} = \frac{24.687,5}{524,7} = 47,05^{**}$$

Das tabelas A9 e A10, apêndice A, dado que $H_{(I,J-1)} \alpha$, então $H_{5;5;0,05} = 16,3$. Há razão para rejeitar a hipótese da homoscedasticidade, para o nível de significância $\alpha = 0,05$, e concluir pela existência de heteroscedasticidade dos dados.

b) Teste de Bartlett

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2$$

$$\chi_{I-1}^2 = 2,3026 \left\{ \left[\sum_{i=1}^I (J_i - 1) \right] \log \hat{\sigma}^2 - \sum_{i=1}^I (J_i - 1) \log \hat{\sigma}_i^2 \right\}$$

C: fator de correção:

$$C = 1 + \frac{1}{3(I-1)} \left[\sum_{i=1}^I \frac{1}{J_i - 1} - \frac{1}{\sum_{i=1}^I (J_i - 1)} \right]$$

Trat.	J_i	$\hat{\sigma}^2$	$\log \hat{\sigma}^2$	$(J_i - 1) \log \hat{\sigma}^2$	$\frac{1}{(J_i - 1)}$
A	6	24.687,50	4,39248	21,9624	1/5
B	6	17.067,90	4,23218	21,1609	1/5
C	6	8.753,10	3,94216	19,7108	1/5
D	6	2.008,4	3,30285	16,5143	1/5
E	6	524,67	2,71988	13,5994	1/5
				92,9478	

Considerando que o número de repetições é o mesmo, então:

$$M = 2,3026 \left[I(J-1) \log \bar{\sigma}^2 - (J-1) \sum_{i=1}^I \log \hat{\sigma}_i^2 \right]$$

$$\chi_{I-1}^2 = 2,3026 (25 \times \log 10.608,314 - 92,9478) = 17,7147 \quad (4 \text{ graus de liberdade}).$$

$$\hat{\sigma}^2 = 10.608,314$$

$$C = 1 + \frac{1}{3 \times 4} \left(1 - \frac{1}{25} \right) = 1,08$$

$$\chi_{I-1}^2 = \frac{17,7147}{1,08} = 16,4025$$

$$\chi_{4;0,05}^2 = 9,49$$

De acordo com o programa SAS, **Prog6**, tem-se:

Teste de Homogeneidade de variâncias de Bartlett			
F.V.	G.L.	Chi-Square (Qui-Quadrado)	Pr > ChiSq
Trat	4	16,4025	0,0025

Considerando um nível de significância $\alpha = 0,05$, conclui-se que não há razão para aceitar a hipótese de homoscedasticidade (H_0 : homogeneidade de variâncias dos tratamentos). Ou seja, admite-se a existência de heteroscedasticidade.

c) Teste de Cochran

$$C = \frac{\hat{\sigma}_{\max}^2}{\sum_{i=1}^I \hat{\sigma}_i^2} = \frac{24.687,5}{53.038,9} = 0,4655$$

Dado que $C_{I; J-1; 0,05} = 0,5441$ (Tabela A7 do apêndice A), em que: $I = 5$ e $(J - 1) = 5$ graus de liberdade, tal que I é o número de variâncias e J é o número de repetições.

Considerando que $C_{\text{Calculado}} < C_{I; J-1; 0,05}$, então há razão para rejeitar, ao nível de significância de $\alpha = 0,05$, a homogeneidade de variâncias. Isto é, rejeita-se:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2$$

Exercício 2.5.2. Sendo o experimento inteiramente casualizado, apresentado no exercício 2.5.1, verificar, pelo método Box-Cox, a transformação mais adequada.

Tabela 2.5.3. Aplicação do método de Box-Cox.

TRAT	$\hat{m}$	$\hat{\sigma}^2$	$\log \hat{m}$	$\log \hat{\sigma}^2$
A	5.477,70	24.687,5	3,73860	4,39248
B	1.678,70	17.067,9	3,22497	4,23218
C	819,30	8.753,1	2,91344	3,94216
D	194,00	2.008,4	2,28780	3,30285
E	172,67	524,7	2,23722	2,71989

A equação linear simples ajustada: $\log(\hat{\sigma}_i^2) = 0,7838 + 1,0186 \text{ Log}(\hat{m}_i)$.

Tabela 2.5.4. ANOVA da regressão linear simples.

F.V.	G.L.	S.Q.	Q.M.	F	p-valor
Regressão	1	1,6822	1,6822	19,72	0,021
Resíduo	3	0,2559	0,0853		
Total	4	1,9381			

$$\lambda = 1 - \frac{\hat{b}}{2} = 1 - \frac{1,0186}{2} = 0,4907$$

De acordo com a tabela 2.5.1, como $\lambda = 0,4907$ está mais próximo de 0,5, recomenda-se usar $\sqrt{Y}$ ou $\sqrt{X}$, ao invés de $\log Y$ ou $\log X$, na análise do ensaio apresentado no exercício 2.5.1.

2.6. ANÁLISE DE UM ENSAIO INTEIRAMENTE AO ACASO COM K AMOSTRAS POR PARCELA

Exercício 2.6.1. Sejam considerados os dados de um ensaio inteiramente ao acaso com K amostras por parcelas (modelo fixo), considerando quatro espécies, E1, E2, E3 e E4, cultivadas em soluções com nutrientes, em que foi avaliado depois de dois meses o crescimento em altura (em cm). Em cada tratamento, foram utilizados três vasos. Em cada vaso foram inseridas quatro plantas P1, P2, P3 e P4, e assim totalizando 12 plantas por tratamento. A Tabela 2.6.1 apresenta os dados simulados de crescimento em centímetros.

Na tabela 2.6.2, estão os componentes de variância do ensaio inteiramente casualizado com K amostras por parcela (modelo fixo), como visto na tabela 2.4.4.

Tabela 2.6.1. Dados de crescimento em centímetros.

	E1			E2			E3			E4		
	V11	V12	V13	V21	V22	V23	V31	V32	V33	V41	V42	V43
P1	8,2	5,8	12,0	11,7	8,2	10,5	11,7	12,9	12,9	19,9	15,2	8,5
P2	9,3	10,4	7,0	12,0	8,2	9,3	10,5	14,0	10,5	18,0	16,4	16,3
P3	7,0	12,9	5,8	9,3	7,0	9,3	15,7	11,7	15,2	21,1	18,7	12,4
P4	10,5	11,7	7,0	8,2	9,3	11,7	10,5	11,7	12,9	19,9	15,2	20,1
	35,0	40,8	31,8	41,2	32,7	40,8	48,4	50,3	51,5	78,9	65,5	57,3
	107,6			114,7			150,2			201,7		

Tabela 2.6.2. Os componentes de variância do ensaio inteiramente casualizado com K amostras por parcela (modelo fixo).

F.V.	G.L.	E(QM)
Tratamentos	$I - 1$	$\sigma_e^2 + K\sigma_e^2 + JK\phi_t$
Entre parcelas d. Trat.(erro experimental)	$I(J-1)$	$\sigma_e^2 + K\sigma_e^2$
Entre amostras d. Parcelas (erro amostral)	$IJK-1$	σ_e^2
Total	$IJK-1$	

$$x_{ijk} = m + t_i + e_{ij} + \varepsilon_{ijk} \quad \phi_t = \frac{\sum_{i=1}^I t_i^2}{I}$$

m : é a média geral;

t_i : é o efeito de tratamentos, em que $i = 1, \dots, I$ (I : número de tratamentos);

e_{ij} : é o erro experimental (Ou erro entre), em que $j = 1, \dots, J$ (J : número de parcelas por tratamento), tal que: $E(e_{ij}^2) = \sigma_e^2$;

ε_{ijk} : é o erro amostral (Ou erro dentro), em que $k = 1, \dots, K$ (K : número de amostras por parcela), tal que $E(\varepsilon_{ijk}^2) = \sigma_e^2$.

a) Soma de quadrados total:

$$C = \frac{(8,2 + 9,3 + \dots + 12,4 + 20,1)^2}{48} = \frac{574,2^2}{48} = 6.868,8675$$

$$SQ_{TOTAL} = 8,2^2 + 9,3^2 + \dots + 12,4^2 + 20,1^2 - C$$

$$SQ_{TOTAL} = 7610,28 - 6.868,8675 = 741,4125$$

b) Soma de quadrados de tratamentos:

$$SQ_{TRAT} = \frac{1}{12} (107,6^2 + 114,7^2 + 150,2^2 + 201,7^2) - 6.868,8675 = 462,5308$$

c) Soma de quadrados entre vasos dentro de tratamentos (Erro experimental):

$$SQ_{VASOS (T1)} = \frac{1}{4} (35,0^2 + 40,8^2 + 31,8^2) - \frac{107,6^2}{12} = 10,4067$$

$$SQ_{VASOS (T2)} = \frac{1}{4} (41,2^2 + 32,7^2 + 40,8^2) - \frac{114,7^2}{12} = 11,5017$$

$$SQ_{VASOS (T3)} = \frac{1}{4} (48,4^2 + 50,3^2 + 51,5^2) - \frac{150,2^2}{12} = 1,2217$$

$$SQ_{VASOS (T4)} = \frac{1}{4} (78,9^2 + 65,5^2 + 57,3^2) - \frac{201,7^2}{12} = 59,4466$$

$$SQ_{VASOS (TRAT)} = SQ_{VASOS (T1)} + SQ_{VASOS (T2)} + SQ_{VASOS (T3)} + SQ_{VASOS (T4)}$$

$$SQ_{VASOS (TRAT)} = 10,4067 + 11,5017 + 1,2217 + 59,4466 = 82,5767$$

Ou, simplesmente:

$$SQ_{VASOS (TRAT)} = SQ_{VASOS} - SQ_{TRAT} = 545,1075 - 462,5308 = 82,5767$$

Em que:

$$SQ_{VASOS} = \frac{1}{4} (35,0^2 + 40,8^2 + \dots + 65,5^2 + 57,3^2) - C$$

$$SQ_{VASOS} = \frac{1}{4} (29.655,9) - C = 545,1075$$

d) Soma de quadrados entre plantas dentro dos vasos dentro de tratamentos (Erro de amostragem):

$$SQ_{ENTRE PLANTAS (P) D. VASOS} = SQ_{P D. V11} + SQ_{P D. V12} + \dots + SQ_{P D. V42} + SQ_{P D. V43}$$

Em que:

$$SQ_{P D, V11} = (8,2^2 + 9,3^2 + 7,0^2 + 10,5^2) - \frac{35,0^2}{4} = 6,73$$

$$SQ_{P D, V12} = (5,8^2 + 10,4^2 + 12,9^2 + 11,7^2) - \frac{40,8^2}{4} = 28,94$$

$$SQ_{P D, V13} = (12,0^2 + 7,0^2 + 5,8^2 + 7,0) - \frac{31,8^2}{4} = 22,83$$

$$SQ_{P D, V21} = (11,7^2 + 12,0^2 + 9,3^2 + 8,2^2) - \frac{41,2^2}{4} = 10,26$$

$$SQ_{P D, V22} = (8,2^2 + 8,2^2 + 7,0^2 + 9,3^2) - \frac{32,7^2}{4} = 2,6475$$

$$SQ_{P D, V23} = (10,5^2 + 9,3^2 + 9,3^2 + 11,7^2) - \frac{40,8^2}{4} = 3,96$$

$$SQ_{P D, V31} = (11,7^2 + 10,5^2 + 15,7^2 + 10,5^2) - \frac{48,4^2}{4} = 18,24$$

$$SQ_{P D, V32} = (12,9^2 + 14,0^2 + 11,7^2 + 11,7^2) - \frac{50,3^2}{4} = 3,6675$$

$$SQ_{P D, V33} = (12,9^2 + 10,5^2 + 15,2^2 + 12,9^2) - \frac{51,5^2}{4} = 11,0475$$

$$SQ_{P D, V41} = (19,9^2 + 18,0^2 + 21,1^2 + 19,9^2) - \frac{78,9^2}{4} = 4,9275$$

$$SQ_{P D, V42} = (15,2^2 + 16,4^2 + 18,7^2 + 15,2^2) - \frac{65,5^2}{4} = 8,1675$$

$$SQ_{P D, V43} = (8,5^2 + 16,3^2 + 12,4^2 + 20,1^2) - \frac{57,3^2}{4} = 74,8875$$

$$SQ_{ENTRE PLANTAS (P) D. VASOS} = 6,73 + 28,94 + \dots + 8,1675 + 74,8875 = 196,3050$$

Ou, simplesmente:

$$SQ_{ENTRE PLANTAS (P) D. VASOS} = SQ_{TOTAL} - SQ_{TRAT} - SQ_{VASOS (TRAT)}$$

$$SQ_{ENTRE PLANTAS (P) D. VASOS} = 741,4125 - 462,5308 - 82,5767 = 196,3050$$

Tabela 2.6.3. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Entre vasos	(11)	(545,1075)		
Tratamentos	3	462,5308	154,1769	14,94**
Entre vasos d. Trat. (erro experimental)	8	82,5767	10,3221	1,89 <i>ns</i>
Entre plantas d. vasos (erro amostral)	36	196,3050	5,4529	
Total	47	741,4125		

$$F_{3; 8; 0,05} = 4,07 \quad F_{8; 36; 0,05} = 2,22$$

Tabela 2.6.4. ANOVA no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Entre vasos	11	545,1075000	49,5552273	9,09	< 0,0001
Tratamentos	3	462,5308333	154,1769444	14,94	0,0012
Entre vasos d. Trat.	8	82,57666667	10,32208333	1,89	0,0916
Entre plantas d. vasos	36	196,3050000	5,4529167		
Total	47	741,4125000			

e) Teste de Tukey para comparar os tratamentos E1, E2, E3, E4:

O capítulo 3 apresenta, com mais detalhes, o método de Tukey (1949b), o qual se baseia na construção de intervalos de confiança para contrastes entre duas médias:

$$Y_{iu} = m_i - m_u \quad (i \neq u)$$

Sendo o coeficiente de confiança $1 - \alpha$ para o conjunto de todos os intervalos entre duas médias de tratamentos, assegurando um nível de significância conjunto α em todos os testes, o intervalo de confiança para Y_{iu} em experimentos balanceados é dado por:

$$[(\hat{m}_i - \hat{m}_u) \pm q_{(I, n-I)} \sqrt{\frac{QM_{RES}}{J}}]$$

A distribuição de $q(I, n - I)$ é denominada “Studentized Range”, com parâmetros I (Número de tratamentos) e $n - I$ (Número de graus de liberdade do resíduo). Da tabela A14, no apêndice A, tem-se: $q_{4; 8; 0,05} = 4,53$.

$$\Delta = q_{4; 8; 0,05} \times \sqrt{\frac{QM_{RES}}{J}} = 4,53 \sqrt{\frac{10,3221}{12}} = 4,20$$

Trat.	Médias	
E4	16,81	a
E3	12,52	b
E2	9,56	b
E1	8,97	b

Médias seguidas pela mesma letra não diferem entre si, ao nível de 5% de significância, pelo teste de Tukey. A espécie E4 apresentou resultado significativo em relação às espécies E3, E2 e E1, que apresentaram resultados não significantes entre si. Recomenda-se cultivar a espécie E4.

f) Teste de Tukey, no SAS, para comparar os tratamentos E1, E2, E3, E4:

O programa SAS (**Prog8**), em anexo no apêndice B, foi usado na análise do ensaio inteiramente ao acaso com K amostras por parcela.

Alfa (α)	0,05
Graus de liberdade do resíduo	8
Quadrado médio do resíduo	10,32208
Amplitude total estudentizado (Valor crítico)	4,52880
Diferença mínima de significância	4,2003

Tukey (Agrupamento)	Média	N	Trat.
A	16,808	12	E4
B	12,517	12	E3
B	9,558	12	E2
B	8,967	12	E1

2.7. ANÁLISE DE COVARIÂNCIA

A análise de covariância (ANACOVA) engloba conjuntamente as concepções das análises de variância e de regressão. A ANACOVA utiliza uma ou várias variáveis auxiliares (Covariáveis) para melhor interpretar os dados relativos à variável de interesse. A aplicação da ANACOVA pode, portanto, ser usada como mecanismo de controle local, assim como para melhorá-lo, ou seja, permite que se faça um “ajuste” do efeito de uma variável resposta que sofreu influência de outra variável ou de uma causa de variação não controlada.

Em alguns experimentos pode ser difícil ou até impossível se obter unidades experimentais homogêneas. Por exemplo, sementes de mesmo peso, em um experimento com a espécie *Carapa guianensis* (Andiroba), podem não ser obtidas. Da mesma forma, pode-se não obter, em ensaios na área de tecnologia da madeira, corpos de prova com mesmo tamanho. Entretanto, em todos esses casos, quando a natureza da unidade experimental é conhecida e pode ser mensurada — ainda que essa condição inicial (Uma covariável) tenha influência sobre a variável resposta —, é factível fazer uso da técnica de análise de covariância a fim de se corrigir a variável resposta.

A covariável complementa o controle local e, na grande maioria das situações, simplesmente o substitui. A variável medida, na condição inicial da unidade experimental, é chamada de covariável ou variável auxiliar. Obviamente, a covariável necessita estar correlacionada com a variável resposta para que se possa fazer uso da análise de covariância. Em um mesmo experimento, pode haver mais de uma covariável.

Quando existir heterogeneidade entre as unidades experimentais, ou mesmo dificuldades ou restrições na organização dos blocos, o uso da ANACOVA pode melhorar a precisão dos resultados do teste de significância. Por outro lado, em alguns casos, algumas variáveis contínuas podem estar

causando heterogeneidade entre as unidades experimentais. Essas fontes de variabilidade externas são chamadas de “covariáveis”. Logo, um modelo ANACOVA é formado pela inclusão de uma covariável contínua em um delineamento de ANOVA. Entretanto, como a covariável contínua é inserida no modelo como uma variável de uma regressão, uma ANACOVA requer algumas etapas adicionais que devem ser combinadas com o procedimento ANOVA.

A análise de covariância pode funcionar como um controle do erro experimental, portanto aumentando a precisão do experimento. É possível também fazer o ajuste das médias dos tratamentos em função da(s) covariável (eis) e, em alguns casos, estimar observações perdidas durante o experimento.

Em uma situação hipotética, para melhor ilustrar a importância que a covariável tem na ANACOVA, um pesquisador deseja comparar cinco tipos de intervenções silviculturais no manejo de uma determinada floresta natural. Então, uma amostra aleatória de cinco parcelas é implantada para cada tipo de tratamento. Depois de completado o ciclo de corte, o volume de biomassa é mensurado e, nas vinte e cinco unidades experimentais, é realizada uma simples análise de variância. É razoável supor que o número de árvores por unidade experimental exerça influência no volume de biomassa. Assim sendo, a variável número de indivíduos deve ser incluída a fim de melhorar a precisão dos resultados no modelo de análise. O gráfico, considerando as duas variáveis, biomassa e número de plantas, indicará o tipo de tendência, se linear ou não linear.

Por outro lado, é necessário confirmar a existência de uma relação de regressão significativa da variável resposta com a covariável. Caso contrário, incluir a covariável no modelo não melhorará a estimativa das médias dos tratamentos.

Para que uma covariável possa ser considerada na análise, deve ser garantido que ela não seja afetada pelo tratamento. Por exemplo, ao se utilizar como covariável o número de árvores em uma parcela num experimento florestal, deve-se ter a garantia de não haver efeito do tratamento no estoque referente ao número de indivíduos. No caso da variável escolhida como covariável ser influenciada pelos tratamentos, o seu uso é incorreto.

Steel e Torrie (1960) citam os mais importantes usos da análise de covariância:

- a) Auxiliar a interpretação dos dados, principalmente no que diz respeito à natureza dos efeitos dos tratamentos;
- b) Desdobrar uma covariância total, ou soma de produtos cruzados, em partes denominadas componentes;
- c) Controlar o erro e aumentar a precisão;
- d) Ajustar as médias de tratamentos da variável dependente para as diferenças nos conjuntos de valores das variáveis independentes correspondentes;
- e) Estimar dados perdidos.

Neste livro, será abordada a ANACOVA para o modelo de ANOVA delineado para incluir o efeito linear de apenas uma covariável. Entretanto, o modelo ANACOVA pode ser generalizado para incluir também os efeitos quadráticos e cúbicos da covariável.

2.7.1. Análise de covariância para o ensaio inteiramente ao acaso.

Seja considerado o modelo estatístico em um experimento com um fator e uma covariável.

$$Y_{ij} = m + \alpha_i + \beta(X_{ij} - \bar{X}) + e_{ij}$$

m : Valor constante;

α_i : Efeito do i -ésimo tratamento;

β : Coeficiente de regressão linear, tal que $\beta \neq 0$, entre a variável resposta (Y) e a covariável (X);

X_{ij} : valor observado da covariável inerente à j -ésima repetição do i -ésimo tratamento.

Tabela 2.7.1.1. ANACOVA para o ensaio inteiramente ao acaso.

		Soma de quadrados e produtos			Ajuste pela regressão de Y para X		
F.V.	G.L.	y^2	xy	x^2	G.L.	S.Q.	Q.M.
Trat.	$I-1$	T_{yy}	T_{xy}	T_{xx}			
Res.	$I(J-1)$	E_{yy}	E_{xy}	E_{xx}	$I(J-1)-1$	$SQ_{RES(*)} = E_{yy} - \frac{(E_{xy})^2}{E_{xx}}$	$\frac{SQ_{RES(*)}}{I(J-1)-1}$
Total	$IJ-1$	S_{yy}	S_{xy}	S_{xx}	$IJ-2$	$SQ_{(TRAT+RES)(*)} = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$	$\frac{SQ_{(TRAT+RES)(*)}}{IJ-2}$
Trat. Ajust.					$I-1$	$SQ_{(TRAT+RES)(*)} - SQ_{RES(*)}$	$\frac{SQ_{(TRAT+RES)(*)} - SQ_{RES(*)}}{I-1}$

$$S_{yy} = \sum_{i=1}^I \sum_{j=1}^J Y_{ij}^2 - \frac{Y_{..}^2}{IJ} \quad Y_{..} = \sum_{i=1}^I \sum_{j=1}^J Y_{ij}$$

$$S_{xx} = \sum_{i=1}^I \sum_{j=1}^J X_{ij}^2 - \frac{X_{..}^2}{IJ} \quad X_{..} = \sum_{i=1}^I \sum_{j=1}^J X_{ij}$$

$$S_{xy} = \sum_{i=1}^I \sum_{j=1}^J X_{ij} Y_{ij} - \frac{(X_{..})(Y_{..})}{IJ}$$

$$T_{yy} = \frac{1}{J} \sum_{i=1}^I Y_{i.}^2 - \frac{Y_{..}^2}{IJ} \quad Y_{i.} = \sum_{j=1}^J Y_{ij}$$

$$T_{xx} = \frac{1}{J} \sum_{i=1}^I X_{i.}^2 - \frac{X_{..}^2}{IJ} \quad X_{i.} = \sum_{j=1}^J X_{ij}$$

$$T_{xy} = \frac{1}{J} \sum_{i=1}^I X_i Y_i - \frac{(X_{..})(Y_{..})}{IJ}$$

$$E_{yy} = S_{yy} - T_{yy} \quad E_{xx} = S_{xx} - T_{xx} \quad E_{xy} = S_{xy} - T_{xy}$$

Teste para igualdade de médias dos tratamentos:

A soma de quadrados dos tratamentos, ajustada de acordo com a regressão, é obtida aplicando o método do resíduo condicional, no qual se adiciona às somas de produtos do resíduo as somas correspondentes aos tratamentos. O uso do asterisco indica que foi feito o ajuste da regressão.

$$SQ_{(TRAT+RES)(*)} = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$$

$$SQ_{RES(*)} = E_{yy} - \frac{(E_{xy})^2}{E_{xx}}$$

A soma de quadrados (ajustada) para tratamentos é dada por:

$$SQ_{TRAT(*)} = SQ_{(TRAT+RES)(*)} - SQ_{RES(*)}$$

$$F = \frac{\frac{SQ_{(TRAT+RES)(*)} - SQ_{RES(*)}}{I-1}}{\frac{SQ_{RES(*)}}{I(J-1)-1}}$$

Teste da hipótese de nulidade do coeficiente $\hat{\beta}$, ou seja, $H_0: \beta = 0$, tal que: $\hat{\beta} = \frac{E_{xy}}{E_{xx}}$.

$$F = \frac{\frac{(E_{xy})^2}{(E_{xx})}}{\frac{E_{yy} - \frac{(E_{xy})^2}{E_{xx}}}{I(J-1)-1}} \sim F_{1; I(J-1)-1; \alpha}$$

O ajuste para as observações concernentes às variáveis respostas (Y) pode ser expresso por:

$$Y_{ij} - \beta(X_{ij} - \bar{X}) = m + \alpha_i + e_{ij}$$

O ajuste para as médias pode ser obtido por: $\bar{Y}'_i = \bar{Y}_i - \hat{\beta}(\bar{X}_i - \bar{X})$.

Exercício 2.7.1.1. Em um experimento no manejo de uma determinada área florestal natural (Dados simulados), no qual um pesquisador deseja comparar cinco tipos de intervenções silviculturais (M1, M2, M3, M4 e M5), uma amostra aleatória de cinco parcelas é implantada para cada tipo de tratamento e, depois de completado o ciclo de corte (Rotação), o volume (m³ por parcela) de biomassa (Y) e o número de árvores (X) são mensurados nas vinte e cinco unidades experimentais. Considerando a descrição desse experimento, efetuar a análise de covariância simples.

Tabela 2.7.1.2. Volume de madeira (m³) e número de árvores por parcela.

	M1		M2		M3		M4		M5	
	Y	X	Y	X	Y	X	Y	X	Y	X
1	66,9	30	40,7	23	48,7	27	46,9	22	55,4	29
2	68,2	32	54,5	25	56,9	36	49,9	24	57,8	32
3	66,0	35	35,8	16	48,4	23	45,3	21	42,7	21
4	58,1	30	28,7	19	40,8	19	43,8	23	36,7	14
5	57,7	26	15,0	13	32,4	10	45,6	20	46,7	21
Total	316,9	153	174,7	96	227,2	115	231,5	110	239,3	117

a) Análise de variância, considerando a variável biomassa (Y : volume de madeira por parcela).

A tabela 2.7.1.3 apresenta a análise de variância para a variável biomassa (Y).

Tabela 2.7.1.3. ANOVA para a variável biomassa (Y).

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	4	2.078,5296	519,6324	6,39
Resíduo	20	1.627,2040	81,3602	
Total	24	3.705,7336		

$$F_{4; 20; 0,05} = 2,87$$

Como $F > F_{4; 20; 0,05}$, então há razão para rejeitar H_0 :

$$H_0 : m_1 = m_2 = m_3 = m_4 = m_5$$

b) Análise de variância para a covariável número de árvores (X : número de árvores por parcela).

A tabela 2.7.1.4 apresenta a análise de variância da covariável número de árvores (X).

Tabela 2.7.1.4. ANOVA para a variável número de árvores (X).

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	4	356,5600	89,1400	2,46 ns
Erro	20	725,2000	36,2600	
Total	24	1.081,7600		

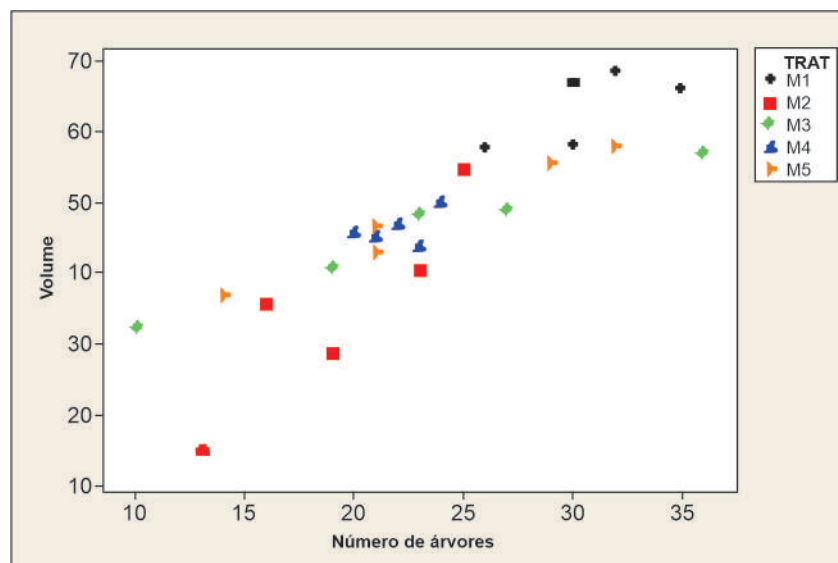
Como $F < F_{4; 20; 0,05}$, então há razão para não rejeitar H_0 :

$$H_0 : m_1 = m_2 = m_3 = m_4 = m_5$$

Não houve influência dos tratamentos na covariável. Logo não há restrição para analisar o experimento usando a covariável número de árvores (X).

c) Análise de covariância

A figura 2.7.1.1 mostra a existência de uma tendência linear entre a variável resposta volume e a variável número de árvores por parcela.



A figura 2.7.1.1. Tendência entre a variável resposta e a covariável.

c1) Cálculo das somas de quadrados total, considerando Y , X e XY :

$$S_{yy} = \sum_{i=1}^I \sum_{j=1}^J Y_{ij}^2 - \frac{Y_{..}^2}{IJ} = 60.311,66 - \frac{1189,6^2}{25} = 3.705,7336$$

$$S_{xx} = \sum_{i=1}^I \sum_{j=1}^J X_{ij}^2 - \frac{X_{..}^2}{IJ} = 15.053,0 - \frac{591,0^2}{25} = 1.081,76$$

$$S_{xy} = \sum_{i=1}^I \sum_{j=1}^J X_{ij} Y_{ij} - \frac{(X_{..})(Y_{..})}{IJ} = 29.877,5 - \frac{591 \times 1.189,6}{25} = 1.755,3560$$

c2) Cálculo das somas de quadrados de tratamentos, considerando Y , X e XY :

$$T_{yy} = \frac{1}{J} \sum_{i=1}^I Y_{i.}^2 - \frac{Y_{..}^2}{IJ} \quad Y_{i.} = \sum_{j=1}^J Y_{ij}$$

$$T_{yy} = \frac{1}{5} (316,9^2 + 174,7^2 + 227,2^2 + 231,5^2 + 239,3^2) - \frac{1.189,6^2}{25} = 2.078,5296$$

$$T_{xx} = \frac{1}{J} \sum_{i=1}^I X_{i.}^2 - \frac{X_{..}^2}{IJ} \quad X_{i.} = \sum_{j=1}^J X_{ij}$$

$$T_{xx} = \frac{1}{5} (153^2 + 96^2 + 115^2 + 110^2 + 117^2) - \frac{591^2}{25} = 356,56$$

$$T_{xy} = \frac{1}{5} (316,9 \times 153 + 174,7 \times 96 + 227,2 \times 115 + 231,5 \times 110 + 239,3 \times 117) - \frac{591 \times 1.189,6}{25}$$

$$T_{xy} = 847,4560$$

c3) Cálculo das somas de quadrados de resíduos, considerando Y , X e XY :

$$E_{yy} = S_{yy} - T_{yy} = 3.705,7336 - 2.078,5296 = 1.627,2040$$

$$E_{xx} = S_{xx} - T_{xx} = 1.081,76 - 356,56 = 725,20$$

$$E_{xy} = S_{xy} - T_{xy} = 1.755,3560 - 847,456 = 907,9$$

Tabela 2.7.1.5. ANACOVA preliminar para o ensaio inteiramente ao acaso.

F.V.	G.L.	Soma de quadrados e produtos		
		y^2	xy	x^2
Tratamentos	4	2078,5296	847,4560	356,56
Resíduo	20	1.627,2040	907,9000	725,20
Total	24	3.705,7336	1.755,3560	1.081,76

c4) Cálculo do coeficiente da regressão ($\hat{\beta}$) e soma de quadrados da regressão linear

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}} = \frac{907,9}{725,2} = 1,2519$$

$$SQ_{REG} = \frac{E_{xy}^2}{E_{xx}} = \frac{(907,9)^2}{725,2} = 1.136,6277$$

Então, a soma de quadrados do resíduo ajustada para a regressão, será:

$$SQ_{RES(*)} = E_{yy} - SQ_{REG} = 1.627,2040 - 1.136,6277 = 490,5763$$

O número de graus de liberdade será 19, pois $I(J-1)-1=5(5-1)-1=19$, haja vista a perda de um grau correspondente à regressão.

Então o quadrado médio residual será obtido como segue:

$$QM_{RES(*)} = \frac{490,5763}{19} = 25,8198$$

c5) Cálculo da soma de quadrados dos resíduos considerando Y ajustado para X :

A soma de quadrados dos tratamentos ajustada, de acordo com a regressão, é obtida aplicando-se o método do resíduo condicional, em que se adiciona às somas de produtos dos resíduos às somas correspondentes aos tratamentos. O uso do asterisco indica que um ajuste foi efetuado para a regressão.

	y^2	xy	x^2
Tratamentos +Resíduo (*)	3.705,7336	1.755,3560	1.081,76

$$SQ_{(TRAT+RES)(*)} = S_{yy} - \frac{(S_{xy})^2}{S_{xx}} = 3.705,7336 - \frac{1.755,3560^2}{1.081,76} = 857,3433$$

$$SQ_{RES(*)} = E_{yy} - \frac{(E_{xy})^2}{E_{xx}} = 1.627,2040 - \frac{907,9^2}{725,20} = 490,5763$$

A soma de quadrados ajustada para tratamentos com quatro graus de liberdade, é obtida como segue:

$$SQ_{TRAT(Aj.)} = SQ_{(TRAT+RES)(*)} - SQ_{RES(*)} = 857,3433 - 490,5763 = 366,7670$$

c6) Teste para igualdade de médias de tratamentos:

$$F = \frac{\frac{SQ_{(TRAT+RES)(*)} - SQ_{RES(*)}}{I-1}}{\frac{SQ_{RES(*)}}{I(J-1)-1}} = \frac{\frac{366,7670}{4}}{\frac{490,5763}{5(4)-1}} = \frac{91,6917}{25,8198} = 3,55$$

A Tabela 2.7.1.6 apresenta a análise de covariância considerando o volume da biomassa versus o número de árvores.

Tabela 2.7.1.6. ANACOVA: volume da biomassa vs número de árvores.

F.V.	G.L.	Soma de quadrados e produtos			Ajuste pela regressão		
		SQ_Y	SPC_{XY}	SQ_X	Y	G.L.	Q.M.
Tratamentos	4	2.078,5296	847,456	356,56			
Resíduo	20	1.627,2040	908,394	725,20	490,5763	19	25,8198
Total	24	3.705,7336	1.755,856	1.081,76	857,3433	23	
Trat (Aj.)					366,7670	4	91,6917

$$F = \frac{QM_{TRAT(Aj.)}}{QM_{RES*}} = \frac{91,6917}{25,8198} = 3,55 \quad F_{4;19;0,05} = 2,90$$

Conclui-se, considerando os tratamentos ajustados e o nível de significância $\alpha = 0,05$, pela rejeição de $H_0 : m_1 = m_2 = \dots = m_5$.

Teste da hipótese $H_0 : \beta = 0$:

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}} = \frac{907,9}{725,20} = 1,2519$$

Logo:

$$F = \frac{\frac{(E_{xy})^2}{(E_{xx})}}{\frac{E_{yy} - \frac{(E_{xy})^2}{E_{xx}}}{I(J-1)-1}} = \frac{\frac{907,9^2}{725,20}}{\frac{1}{19}(1.627,2040 - \frac{907,9^2}{725,20})} = \frac{1.136,6277}{25,8198} = 44,02$$

Dado que $F_{1;19;0,05} = 4,38$ e considerando um nível de significância de $\alpha = 0,05$, conclui-se pela rejeição de $H_0 : \beta = 0$.

Tabela 2.7.1.7. ANACOVA: volume vs número de árvores no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
TRAT	4	366,767007	91,691752	3,55	0,0252
X	1	1136,627703	1136,627703	44,02	< 0,0001

c7) Cálculo das médias ajustadas para os tratamentos ($\bar{Y}'_i$):

$$\bar{Y}'_i = \bar{Y}_i - \hat{\beta}(\bar{X}_i - \bar{X}) \quad \bar{X} = \frac{\sum_{i=1}^I \sum_{j=1}^J X_{ij}}{IJ} = \frac{591}{25} = 23,64$$

	$\bar{Y}'_i = \bar{Y}_i - 1,2519(\bar{X}_i - 23,64)$		
	$\bar{X}_i$	$\bar{Y}_i$	$\bar{Y}'_i$
1	30,6	63,38	54,6668
2	19,2	34,94	40,4984
3	23,0	45,44	46,2412
4	22,0	46,30	48,3531
5	23,4	47,86	48,1605

Para efetuar comparações entre duas médias de tratamentos ajustadas, a estimativa da variância do contraste é dada pela seguinte expressão:

$$\hat{V}(\hat{Z}) = \sigma^2 \left[\frac{2}{J} + \frac{(\bar{X}_i - \bar{X}_u)^2}{E_{xx}} \right]$$

$$\hat{V}(\hat{Z}) = 25,8198 \left[\frac{2}{5} + \frac{(\bar{X}_i - \bar{X}_u)^2}{725,2} \right]$$

$$\hat{V}(\hat{Z}) = 25,8198 \left[0,4 + \frac{(\bar{X} - \bar{X}_u)^2}{725,2} \right]$$

Na comparação entre os tratamentos 1 e 2 pelo teste de Tukey, tem-se:

$$\hat{V}(\hat{Z}) = 25,8198 \left[0,4 + \frac{(\bar{X}_1 - \bar{X}_2)^2}{725,2} \right]$$

$$\hat{V}(\hat{Z}) = 25,8198 \left[0,4 + \frac{(30,6 - 19,2)^2}{725,2} \right] = 14,954976$$

$$\Delta = q_{5; 19; 0,05} \sqrt{\frac{1}{2} V(\hat{Z})} = 4,25283 \sqrt{\frac{1}{2} 14,954976} = 11,629362$$

Em que: Δ é a diferença mínima significativa.

$$\bar{Y}'_1 - \bar{Y}'_2 = 54,6668 - 40,4984 = 14,1684 > \Delta$$

Trat.	$\bar{Y}'_i$	
1	54,6668	a
2	40,4984	b

Escrevendo na forma de intervalo de confiança, tem-se:

$$IC : (14,1684 - 11,6294 \leq m_1 - m_2 \leq 14,1684 + 11,6294) = 0,95$$

$$IC : (2,539 \leq m_1 - m_2 \leq 25,797) = 0,95$$

Como $\bar{Y}'_1 - \bar{Y}'_2 = 54,6668 - 40,4984 = 14,1684 > \Delta$, e o valor zero está fora do intervalo de confiança, conclui-se que há razão para rejeitar $H_0 : m_1 = m_2$.

Na comparação entre os tratamentos 4 e 5, pelo teste de Tukey, tem-se:

$$\hat{V}(\hat{Z}) = 25,8198 \left[0,4 + \frac{(\bar{X}_4 - \bar{X}_5)^2}{725,2} \right]$$

$$\hat{V}(\hat{Z}) = \frac{490,5763}{19} \left[0,4 + \frac{(22,0 - 23,4)^2}{725,2} \right] = 10,397703$$

$$\Delta = q_{5;19;0,05} \sqrt{\frac{1}{2} V(\hat{Z})} = 4,25283 \sqrt{\frac{1}{2} (10,397703)} = 9,696873$$

$$\bar{Y}'_4 - \bar{Y}'_5 = 48,3531 - 48,1605 = 0,1926 < \Delta$$

Trat.	$\bar{Y}'_i$	
4	48,3531	a
5	48,1605	a

Escrevendo na forma de intervalo de confiança, tem-se:

$$IC : (0,1926 - 9,6969 \geq m_4 - m_5 \geq 0,1926 + 9,6969) = 0,95$$

$$IC : (-9,504 \geq m_4 - m_5 \geq 9,8895) = 0,95$$

Como $\bar{Y}'_4 - \bar{Y}'_5 = 48,3531 - 48,1605 = 0,1926 < \Delta$, assim como ocorre o valor zero dentro do intervalo de confiança, conclui-se que há razão para aceitar $H_0 : m_4 = m_5$.

A tabela 2.7.1.8 mostra os resultados da aplicação do teste de Tukey, na forma de intervalo de confiança, realizado no SAS pelo programa **Prog9**.

Tabela 2.7.1.8. Teste de Tukey realizado pelo SAS.

i	j	Diferença entre as médias	Intervalo de Confiança (1- α)	
1	2	14,167992	2,538623	25,797361
1	3	8,425328	-2,157472	19,008128
1	4	6,313398	-4,513035	17,139830
1	5	6,506100	-3,986257	16,998458
2	3	-5,742664	-15,644569	4,159241
2	4	-7,854595	-17,648607	1,939418
2	5	-7,661892	-17,615687	2,291903
3	4	-2,111931	-11,792859	7,568998
3	5	-1,919228	-11,586178	7,747722
4	5	0,192703	-9,504177	9,889582

O ajuste das médias da variável resposta de acordo com os valores da covariável só se justifica, em geral, se as diferenças das médias da covariável não forem devidas aos tratamentos aplicados. Logo, considerando que a análise de variância correspondente à variável número de árvores foi não significativa ($F = 2,46ns$), denotando não haver significância estatística entre os tratamentos, então se pode usar uma estimativa média para a variância $V(\hat{Z})$ aplicável a qualquer contraste entre dois tratamentos, por meio da expressão sugerida por Finney (1946):

$$\hat{V}(\hat{Z}) = \frac{2\hat{\sigma}^2}{J} \left[1 + \frac{QM_{TRAT(X)}}{E_{xx}} \right]$$

Da análise de variância da variável número de árvores, tem-se: $QM_{RES(X)} = 36,26$.

Da análise de variância da variável volume ajustada para número de árvores, tem-se: $\hat{\sigma}^2 = 25,8198$, assim como: $E_{xx} = S_{xx} - T_{xx} = 1.081,76 - 356,56 = 725,20$. Logo:

$$\hat{V}(\hat{Z}) = \frac{2 \times 25,819805}{5} \left[1 + \frac{36,26}{725,2} \right] = 10,844318$$

$$\Delta = q_{5; 19; 0,05} \sqrt{\frac{1}{2} V(\hat{Z})} = 4,25283 \sqrt{\frac{1}{2} 10,844318} = 9,902940$$

Efetuada as comparações entre os pares de médias, tem-se:

Trat.	$\bar{y}'_L$	
1	54,6668	a
4	48,3531	ab
5	48,1605	ab
3	46,2412	ab
2	40,4984	b

Médias seguidas pela mesma letra, ao nível de significância de $\alpha = 0,05$, não diferem pelo teste de Tukey.

Escrevendo na forma de intervalo de confiança, por exemplo, considerando as médias m_1 e m_2 , tem-se:

$$IC : (14,1684 - 9,902940 \geq m_1 - m_2 \geq 14,1684 + 9,902940) = 0,95$$

$$IC : (4,2655 \geq m_1 - m_2 \geq 24,0713) = 0,95$$

2.8. BIBLIOGRAFIA

BARBIN, D. **Componentes de variância. Teoria e aplicações**. 3ª Ed. Piracicaba/SP. FEALQ, 2019. 144 p.

BOX, G.E.P; COX, D. R. An analysis of transformation. **Journal of the Royal Statistical Society**, B, London, v.26, p. 211-243, 1964.

COCHRAN, W. G. The distribution of quadratic forms in a normal system, with applications to analysis of covariance. **Mathematical proceedings of the Cambridge Philosophical Society**, 30 (2): 178 – 191, 1934.

FERREIRA, D. F. **Uso de recursos computacionais**. Departamento de ciências exatas. UFLA. Lavras-Minas Gerais, 2007. 169 p.

FINNEY, D. J. Standard errors of yields adjusted for regression on an independent measurement. **Biometrics Bull.**, 2: 53-55, 1946.

HICKS, C. R. **Fundamental concepts in the design of experiments**. 2^o ed. Nova York: Holt, Rinehart and Winston, 1973. 349 p.

STEEL, R. G. D; TORRIE, J. H. **Principles and procedures of statistics: with special reference to the biological sciences**. New York: Mc Graw-Hill Book Company, Inc., 1960. 481 p.

TUKEY, J. W. One degree of freedom for non-additivity. **International Biometrics Society**, 5(3): 232–242, 1949a.

TUKEY, J. W. Comparing individual means in the analysis of variance, **Biometrics**, v. 5, p. 99-114, 1949b.

CAPÍTULO 3

TESTES DE COMPARAÇÕES MÚLTIPLAS

Este capítulo apresenta definições, conceitos e aplicações de diferentes testes de comparações múltiplas de contrastes de médias. Define a variância e a covariância entre dois contrastes de médias dos tratamentos. Trata do intervalo de confiança de contrastes. Aborda contrastes ortogonais e apresenta o método dos contrastes ortogonais com aplicação dos testes “t” e F. Aborda os testes de comparações múltiplas de Tukey, Scheffé, Duncan, SNK (Student, Newman, Keuls), Dunnett e o de Bonferroni. Efetua as comparações entre os testes de Tukey, Scheffé, Bonferroni e Duncan. Discute ainda o cálculo do número ideal de repetições em um experimento casualizado em blocos.

3.1 CONTRASTES DE MÉDIAS

Uma vez constatada a existência de efeito significativo de um tratamento, ou fator estudado, por meio da análise de variância dos resultados, as diferenças entre as médias dos tratamentos podem ser detectadas por meio de técnicas estatísticas denominadas de comparações múltiplas.

Seja considerada $Y = c_1X_1 + c_2X_2 + \dots + c_IX_I$ uma função linear. Se $\sum_{i=1}^I c_i = 0$, então Y é definido como um contraste das variáveis X , sendo este o exemplo do contraste Y_1 :

$$Y_1 = 3X_1 - X_2 - X_3 - X_4$$

Quando as variáveis correspondem aos valores das médias dos tratamentos, há os denominados contrastes de médias.

$$Y_1 = 3m_1 - m_2 - m_3 - m_4$$

Por outro lado, como as médias verdadeiras não são conhecidas, para construir as estimativas dos contrastes, existe a necessidade de obter suas estimativas. Por exemplo:

$$\hat{Y}_1 = 3\hat{m}_1 - \hat{m}_2 - \hat{m}_3 - \hat{m}_4$$

Pode-se formar um número muito grande de contrastes. Entretanto, deve-se trabalhar com aqueles com maior importância para a pesquisa.

a) Variância de contraste entre médias de tratamentos

Sendo o contraste $Y = c_1m_1 + c_2m_2 + \dots + c_Im_I$ e seu estimador $\hat{Y} = c_1\hat{m}_1 + c_2\hat{m}_2 + \dots + c_I\hat{m}_I$, então se as médias são independentes e o número de repetições J_i , em que $i = 1, \dots, I$, tal que I é o número de tratamentos, tem-se:

$$V(\hat{Y}) = c_1^2 V(\hat{m}_1) + c_2^2 V(\hat{m}_2) + \dots + c_I^2 V(\hat{m}_I)$$

$$V(\hat{Y}) = c_1^2 \frac{\sigma_1^2}{J_1} + c_2^2 \frac{\sigma_2^2}{J_2} + \dots + c_I^2 \frac{\sigma_I^2}{J_I} = \sum_{i=1}^I \frac{c_i^2 \sigma_i^2}{J_i}$$

Se $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2 = \sigma^2$, então:

$$\hat{V}(\hat{Y}) = \left(\frac{c_1^2}{J_1} + \frac{c_2^2}{J_2} + \dots + \frac{c_I^2}{J_I} \right) \hat{\sigma}^2 = \hat{\sigma}^2 \sum_{i=1}^I \frac{c_i^2}{J_i}$$

Por outro lado, se $J_1 = J_2 = \dots = J_I = J$, tem-se:

$$\hat{V}(\hat{Y}) = (c_1^2 + c_2^2 + \dots + c_I^2) \frac{\hat{\sigma}^2}{J} = \frac{\hat{\sigma}^2}{J} \sum_{i=1}^I c_i^2$$

O erro padrão do contraste é obtido por:

$$s(\hat{Y}) = \sqrt{\hat{V}(\hat{Y})}$$

b) Covariância de dois contrastes

Sejam consideradas $\hat{Y}_1$ e $\hat{Y}_2$ as estimativas dos contrastes Y_1 e Y_2 , tal que as estimativas das médias $\hat{m}_1, \hat{m}_2, \dots, \hat{m}_I$ tenham sido obtidas, respectivamente, com $J_1, J_2, \dots, J_I$ repetições:

$$\begin{aligned}\hat{Y}_1 &= a_1 \hat{m}_1 + a_2 \hat{m}_2 + \dots + a_I \hat{m}_I \\ \hat{Y}_2 &= b_1 \hat{m}_1 + b_2 \hat{m}_2 + \dots + b_I \hat{m}_I\end{aligned}$$

O estimador da covariância entre $\hat{Y}_1$ e $\hat{Y}_2$ é definido por:

$$C\hat{o}v(\hat{Y}_1, \hat{Y}_2) = a_1 b_1 \hat{V}(\hat{m}_1) + a_2 b_2 \hat{V}(\hat{m}_2) + \dots + a_I b_I \hat{V}(\hat{m}_I)$$

Em que:

$$\hat{V}(\hat{m}_i) = \frac{\hat{\sigma}_i^2}{J_i}$$

Logo:

$$C\hat{o}v(\hat{Y}_1, \hat{Y}_2) = a_1 b_1 \frac{\hat{\sigma}_1^2}{J_1} + a_2 b_2 \frac{\hat{\sigma}_2^2}{J_2} + \dots + a_I b_I \frac{\hat{\sigma}_I^2}{J_I}$$

Se $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2 = \sigma^2$:

$$C\hat{o}v(\hat{Y}_1, \hat{Y}_2) = \left(\frac{a_1 b_1}{J_1} + \frac{a_2 b_2}{J_2} + \dots + \frac{a_I b_I}{J_I} \right) \sigma^2$$

Considerando um delineamento balanceado ($J_1 = J_2 = \dots = J_I = J$), tem-se:

$$C\hat{o}v(\hat{Y}_1, \hat{Y}_2) = \frac{\hat{\sigma}^2}{J} \sum_{i=1}^I a_i b_i$$

c) Contrastes Ortogonais

Dois contrastes são ortogonais quando a covariância entre eles é nula. A ortogonalidade de contrastes expressa a independência entre eles, que é importante para efetuar as comparações, haja vista a independência das suas variações. Sendo assim:

$$\frac{a_1 b_1}{J_1} \hat{\sigma}_1^2 + \frac{a_2 b_2}{J_2} \hat{\sigma}_2^2 + \dots + \frac{a_I b_I}{J_I} \hat{\sigma}_I^2 = 0$$

No caso de igualdade de variâncias, $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2 = \sigma^2$:

$$\left(\frac{a_1 b_1}{J_1} + \frac{a_2 b_2}{J_2} + \dots + \frac{a_I b_I}{J_I}\right) \hat{\sigma}^2 = 0 \rightarrow \sum_{i=1}^I \frac{a_i b_i}{J_i} = 0$$

Se houver igualdade de repetições: $J_1 = J_2 = \dots = J_I = J$, tem-se:

$$(a_1 b_1 + a_2 b_2 + \dots + a_I b_I) \frac{\hat{\sigma}^2}{J} = 0 \rightarrow \sum_{i=1}^I a_i b_i = 0$$

d) Intervalo de confiança de contrastes

Como $\hat{Y}$ é normalmente distribuído, pois é uma combinação linear de variáveis aleatórias normalmente distribuídas, e como $\frac{SQ_{RES}}{\sigma^2}$ tem uma distribuição χ^2 com $(n - I)$ graus de liberdade ($n = \sum_{i=1}^I J_i$), também é independente de $\hat{m}_1, \hat{m}_2, \dots, \hat{m}_I$, a estatística

$$\frac{\hat{Y} - Y}{\sqrt{QM_{RES} \times \sum_{i=1}^I \frac{c_i^2}{J_i}}} \text{ tem distribuição "t" de Student com } (n-I) \text{ graus de liberdade.}$$

Assim o intervalo de confiança para o contraste Y , para um coeficiente de confiança $1-\alpha$, será:

$$IC : (\hat{Y} \pm t \sqrt{QM_{RES} \sum_{i=1}^I \frac{c_i^2}{J_i}})$$

Em que “ t ” é o valor ao nível de significância $\frac{\alpha}{2}$ da distribuição “ t ” de Student correspondente ao número de graus de liberdade do resíduo.

e) Método dos contrastes ortogonais com aplicação do teste “ t ”

O teste “ t ” é pouco utilizado, haja vista as premissas requeridas para a aplicação de teste “ t ”, assim como as mesmas comparações podem ser realizadas pelo teste F .

Exigências para uso do teste “ t ”:

- e1) É utilizado para testar um grupo de contrastes ortogonais;
- e2) Os contrastes devem ser escolhidos previamente, isto é, na fase de planejamento do experimento;
- e3) O número de contrastes ortogonais deve ser igual ao número de graus de liberdade de tratamentos;
- e4) Quando se aplica o teste “ t ” ou se decompõe os graus de liberdade em uma análise de variância, o nível de significância conjunto é dado por $[1 - (1 - \alpha)^p]$, em que α é o nível de significância individual e p é o número de contrastes ortogonais.

Seja considerado o contraste $Y = c_1 m_1 + c_2 m_2 + \dots + c_I m_I$, tal que $\sum_{i=1}^I c_i = 0$, em que $\hat{Y} = c_1 \hat{m}_1 + c_2 \hat{m}_2 + \dots + c_I \hat{m}_I$ é a sua estimativa, a estatística do teste é:

$$t = \frac{\hat{Y}_i - 0}{\sqrt{\hat{V}(\hat{Y}_i)}}$$

Em alguns casos pode haver o interesse de confrontar o contraste com um valor arbitrário (ϕ), logo:

$$t = \frac{\hat{Y}_i - \phi}{\sqrt{\hat{V}(\hat{Y}_i)}}$$

f) Método dos contrastes ortogonais com aplicação do teste F

O método dos contrastes ortogonais consiste em definir “a priori” um conjunto de $I-1$ contrastes ortogonais, possibilitando particionar a soma de quadrados dos tratamentos em $I-1$ componentes, tal que cada um possui um grau de liberdade:

$$SQ(\hat{Y}_h) = \frac{(\hat{Y}_h)^2}{\sum_{i=1}^I \frac{c_{ih}^2}{J_i}}, \text{ tal que } SQ_{TRAT} = \sum_{h=1}^{I-1} SQ(\hat{Y}_h)$$

Dada a hipótese $H_{0h} : Y_h = 0$ vs $H_{1h} : Y_h \neq 0$:

A estatística $\frac{SQ(\hat{Y}_h)}{\sigma^2}$ tem sob H_{0h} distribuição “ χ^2 ” com 1 (um) grau de liberdade, sendo independente da $\frac{SQ_{RES}}{\sigma^2}$. Como $\frac{SQ_{RES}}{\sigma^2}$ tem distribuição “ χ^2 ” com $(n-I)$ graus de liberdade, a estatística $F_h = \frac{SQ(\hat{Y}_h)}{QM_{RES}}$ sob H_{0h} tem distribuição F de Snedecor com 1 (um) e $(n-I)$ graus de liberdade.

Na construção dos contrastes a serem testados, as seguintes considerações devem ser observadas:

- i) Três ou mais contrastes serão ortogonais entre si se eles forem ortogonais dois a dois;
- ii) Havendo um experimento com I tratamentos, é possível estabelecer vários grupos de contrastes ortogonais entre si. Entretanto cada grupo terá apenas um conjunto de $(I-1)$ contrastes ortogonais;
- iii) Para assegurar o nível de significância α em cada teste, os contrastes devem ser estabelecidos antes da realização do experimento.

É mais simples construir contrastes de totais de tratamentos do que com as suas médias. Considerando que os tratamentos têm o mesmo número de repetições, tem-se:

$$Y = c_1 T_1 + c_2 T_2 + \dots + c_I T_I$$

Resultando na seguinte soma de quadrados para testar as comparações para o contraste Y :

$$SQ_{\hat{Y}} = \frac{\hat{Y}^2}{J \sum_{i=1}^I c_i^2}$$

De acordo com Silva e Silva (2002), do ponto de vista matemático o desdobramento da soma de quadrados dos tratamentos em contrastes ortogonais é, na realidade, a decomposição vetorial dos tratamentos em componentes ortogonais. Do ponto de vista prático, representa um detalhamento na análise e interpretação do experimento, em que a simples classificação das médias não permite ao pesquisador ter uma boa visualização.

Exercício 3.1.1. Sejam considerados os dados adaptados do trabalho “Aplicação da Vermiculita em Alfobres”, citado em Banzatto e Kronka (1989), efetuado em um delineamento inteiramente casualizado com cinco tratamentos e quatro repetições, visando analisar o crescimento de mudas (em cm) para 60 dias após a semeadura de *Pinus oocarpa*. Testar os contrastes abaixo especificados pelos testes “ t ” e “ F ”.

Tabela 3.1.1. Crescimento das mudas de *Pinus oocarpa* (cm).

Tratamentos	Repetições				Total	Médias
	1	2	3	4		
SC (1)	4,6	5,1	5,8	5,5	21,0	5,250
SC+E (2)	6,0	7,1	7,2	6,8	27,1	6,775
SC+E+NPK (3)	5,8	7,2	6,9	6,7	26,6	6,650
SC+V (4)	5,6	4,9	5,9	5,7	22,1	5,525
SC+V+NPK (5)	5,8	6,4	6,6	6,8	25,6	6,400

T_1 = SC: solo de cerrado

T_2 = SC+E: solo de cerrado + esterco

T_3 = SC+E+NPK: solo de cerrado + esterco + NPK

T_4 = SC+V: solo de cerrado + vermiculita

T_5 = SC+V+NPK: solo de cerrado + vermiculita + NPK

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	4	7,5970	1,8992	7,28 **
Resíduo	15	3,9150	0,2610	
Total	19	11,5120		

$$F_{4; 15; 0,01} = 4,89$$

a) Análise pelo método dos contrastes ortogonais com aplicação do teste “t”

a1) Contraste: solo de cerrado versus os demais tratamentos (Y_1)

$$Y_1 = 4m_1 - m_2 - m_3 - m_4 - m_5 \quad \hat{V}(\hat{Y}) = \frac{\hat{\sigma}^2}{J} \sum_{i=1}^J c_i^2$$

$$\hat{Y}_1 = 4 \times (5,25) - 6,775 - 6,65 - 5,525 - 6,4 = -4,35 \text{ cm}$$

$$\hat{V}(\hat{Y}_1) = \frac{0,2610}{4} [4^2 + (-1)^2 + (-1)^2 + (-1)^2 + (-1)^2] = 1,305$$

$$t = \frac{\hat{Y}_i - 0}{\sqrt{\hat{V}(\hat{Y}_i)}} \rightarrow t = \frac{-4,35}{\sqrt{1,305}} = -3,8079^{**}$$

Dado que $t_{15; 0,05} = 2,13$, há razão para rejeitar, ao nível de significância $\alpha = 0,05$, a hipótese de nulidade: $H_0: Y_1 = 0$.

a2) Contraste: esterco versus vermiculita (Y_2)

$$Y_2 = T_2 + T_3 - T_4 - T_5$$

$$\hat{Y}_2 = 6,775 + 6,65 - 5,525 - 6,4 = 1,5 \text{ cm}$$

$$\hat{V}(\hat{Y}_2) = \frac{0,261}{4} [1^2 + 1^2 + (-1)^2 + (-1)^2] = 0,261$$

$$t = \frac{1,5}{\sqrt{0,261}} = 2,9361^*$$

Dado que $t_{15; 0,05} = 2,13$, há razão para rejeitar, ao nível de significância $\alpha = 0,05$, a hipótese de nulidade: $H_0: Y_2 = 0$.

a3) Contraste: esterco sem NPK versus esterco + NPK (Y_3)

$$Y_3 = T_2 - T_3$$

$$\hat{Y}_3 = 6,775 - 6,65 = 0,125 \text{ cm}$$

$$\hat{V}(\hat{Y}_3) = \left[\frac{(1)^2}{4} + \frac{(-1)^2}{4} \right] \times 0,261 = 0,1305$$

Dado que $t_{15; 0,05} = 2,13$, não há razão para rejeitar, ao nível de significância $\alpha = 0,05$, a hipótese de nulidade: $H_0: Y_3 = 0$.

a4) Contraste: vermiculita sem NPK versus vermiculita + NPK (Y_4)

$$\hat{Y}_4 = T_4 - T_5 \quad \hat{Y}_4 = 5,525 - 6,4 = -0,875$$

$$\hat{V}(\hat{Y}_4) = \left[\frac{(1)^2}{4} + \frac{(-1)^2}{4} \right] \times 0,261 = 0,1305$$

$$t = \frac{-0,875}{\sqrt{0,1305}} = 2,4221^*$$

Dado que $t_{15; 0,05} = 2,13$, há razão para rejeitar, ao nível de significância $\alpha = 0,05$, a hipótese de nulidade: $H_0: Y_4 = 0$.

b) Análise pelo método dos contrastes ortogonais com aplicação do teste F

b1) Contraste: solo de cerrado versus os demais tratamentos (Y_1)

$$Y_1 = 4T_1 - T_2 - T_3 - T_4 - T_5 \text{ (totais dos tratamentos)}$$

$$\hat{Y}_1 = 4 \times (21,0) - 27,1 - 26,6 - 22,1 - 25,6 = -17,4 \text{ cm}$$

$$SQ_{\hat{Y}_1} = \frac{(-17,4)^2}{4[4^2 + (-1)^2 + (-1)^2 + (-1)^2 + (-1)^2]} = 3,7845$$

b2) Contraste: esterco versus vermiculita (Y_2)

$$Y_2 = T_2 + T_3 - T_4 - T_5$$

$$\hat{Y}_2 = 27,1 + 26,6 - 22,1 - 25,6 = 6,0 \text{ cm}$$

$$SQ_{\hat{Y}_2} = \frac{(6,0)^2}{4[1^2 + 1^2 + (-1)^2 + (-1)^2]} = 2,25$$

b3) Contraste: esterco sem NPK versus esterco + NPK (Y_3)

$$Y_3 = T_2 - T_3 \quad \hat{Y}_3 = 27,1 - 26,6 = 0,5 \text{ cm}$$

$$SQ_{\hat{Y}_3} = \frac{(0,5)^2}{4[1^2 + (-1)^2]} = 0,0312$$

b4) Contraste: vermiculita sem NPK versus vermiculita + NPK (Y_4)

$$\hat{Y}_4 = T_4 - T_5$$

$$\hat{Y}_4 = 22,1 - 25,6 = -3,5$$

$$SQ_{\hat{Y}_4} = \frac{(-3,5)^2}{4[1^2 + (-1)^2]} = 1,5313$$

Tabela 3.1.2. ANOVA.

F.V.	G.L	S.Q.	Q.M.	F
Tratamentos	(4)	7,5970	1,8992	7,28
Cerrado vs demais tratamentos	1	3,7845	3,7845	14,50**
Esterco vs vermiculita	1	2,2500	2,2512	8,62*
Esterco S/NPK vs esterco +NPK	1	0,0312	0,0312	0,12 <i>ns</i>
Vermiculita S/NPK vs vermiculita + NPK	1	1,5313	1,5313	5,87 *
Resíduo	15	3,9150	0,2610	
Total	19	11,5120		

$$F_{1;15;0,05} = 4,54 \quad F_{1;15;0,01} = 8,68$$

ns: não significante;

*: significante ao nível de significância igual a $\alpha = 0,05$;

**: significante ao nível de significância igual a $\alpha = 0,01$.

É importante observar que, no teste F , os valores calculados para os quatro contrastes correspondem, respectivamente, aos valores das estatísticas “ t ” ao quadrado, isto é, $(-3,8079)^2 = 14,50$; $(2,9361)^2 = 8,62$; $(0,3460)^2 = 0,12$ e $(2,4221)^2 = 5,87$. Esses resultados ocorrem devido à relação, $F_{1;n;\alpha} = t_{n;\alpha}^2$, que existe entre as funções densidades de F e “ t ” de Student.

3.2. MÉTODO DE COMPARAÇÃO MÚLTIPLA DE TUKEY

Seja considerado um experimento composto de I tratamentos. O método de Tukey (1949b) baseia-se na construção de intervalos de confiança para todos os $\frac{I(I-1)}{2}$ contrastes entre duas médias, isto é:

$$Y_{iu} = m_i - m_u \quad (i \neq u)$$

Considerando o coeficiente de confiança $1-\alpha$ para o conjunto de todos os intervalos, em que as médias dos tratamentos são comparadas duas a duas — mas assegurando um nível de significância conjunto α em todos os testes —, o intervalo de confiança para Y_{iu} em experimentos balanceados, é dado por:

$$[(\hat{m}_i - \hat{m}_u) \pm q_{(I,n-I)} \sqrt{\frac{QM_{RES}}{J}}] \quad (n=I \times J)$$

O valor q , denominado amplitude total estudentizada (Studentized), é o quantil de ordem $1 - \frac{\alpha}{2}$ da seguinte distribuição da estatística:

$$q(I, n-I) = \frac{\max(\hat{m}_i - \hat{m}_u) - \min(\hat{m}_i - \hat{m}_u)}{\sqrt{\frac{QM_{RES}}{J}}}$$

A distribuição de $q(I, n-I)$ é denominada de amplitude total estudentizada “Studentized Range” com, no caso do ensaio inteiramente ao acaso, I e $n-I$ número de graus de liberdade, em que $n-I$ é o número de graus de liberdade do resíduo.

É possível obter-se, de forma aproximada, os limites de confiança para $Y_{iu} = m_i - m_u$ em ensaios não balanceados, por meio da seguinte estatística:

$$[(\hat{m}_i - \hat{m}_u) \pm \frac{q}{\sqrt{2}} \sqrt{QM_{RES}(\frac{1}{J_i} + \frac{1}{J_u})}]$$

i) O método de Tukey se baseia na distribuição entre a maior e a menor estatística de ordem (“range”) de uma amostra;

ii) Ele é exato quando os dados são balanceados e válidos para todos os contrastes entre duas médias.

Exercício 3.2.1. Considerando o ensaio inteiramente ao acaso balanceado de uma espécie florestal e a produção de madeira em m^3 por parcela no final da rotação economicamente aconselhável — submetida a quatro diferentes tipos de manejo M1, M2, M3 e M4 —, com dados simulados, efetuar o teste de Tukey.

Tabela 3.2.1. Produção de madeira em m^3 por parcela.

Trat.	Repetições					
	1	2	3	4	5	6
M1	34,5	38,3	43,5	41,3	39,5	44,0
M2	45,0	53,3	54,0	51,0	49,6	52,3
M3	43,5	54,0	51,8	50,3	53,4	45,6
M4	42,0	36,8	44,3	42,8	43,0	34,1

$\hat{m}_2 = 50,8667$	$\hat{m}_3 = 49,7667$	$\hat{m}_4 = 40,5000$	$\hat{m}_1 = 40,1833$
-----------------------	-----------------------	-----------------------	-----------------------

Tabela 3.2.2. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	3	600,9346	200,3115	13,72**
Resíduo	20	292,0150	4,6007	
Total	23	892,9496		

$$\Delta = q_{(I; n-I; \alpha)} \sqrt{\frac{QM_{RES}}{J}} = q_{(4; 20; 0; 0,05)} \sqrt{\frac{14,6007}{6}} = 3,9583 \times 1,5599 = 6,1748$$

Ordenam-se inicialmente as médias em ordem decrescente, em que médias seguidas pela mesma letra e com o nível de significância $\alpha = 0,05$, não diferem entre si pelo teste de Tukey.

Logo:

Trat.	Médias	
M2	50,8667	a
M3	49,7667	a
M4	40,5000	b
M1	40,1833	b

$\hat{m}_2 - \hat{m}_3 = 1,1000$	$\hat{m}_2 - \hat{m}_4 = 10,3667$	$\hat{m}_2 - \hat{m}_1 = 10,6834$
$\hat{m}_3 - \hat{m}_4 = 9,2667$	$\hat{m}_3 - \hat{m}_1 = 9,5834$	$\hat{m}_4 - \hat{m}_1 = 0,3167$

Tabela 3.2.3. ANOVA e teste de Tukey no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Modelo	3	600,9345833	200,3115278	13,72	< 0,0001
Resíduo	20	292,0150000	14,6007500		
Total	23	892,9495833			

Alfa (α)	0,05
Graus de liberdade do resíduo	20
Quadrado médio do resíduo	14,60075
Valor crítico da amplitude total estudentizada	3,95829
Diferença mínima de significância	6,1748

Tukey (Agrupamento)	Média	N	Trat.
A	50,867	6	M2
A	49,767	6	M3
B	40,500	6	M4
B	40,183	6	M1

A tabela 3.2.3 apresenta a análise realizada pelo programa SAS, **prog10**, em anexo no apêndice B, considerando, para o teste de Tukey, o nível de significância $\alpha = 0,05$.

Exercício 3.2.2. Considerando o ensaio inteiramente ao acaso desbalanceado e a produção de madeira (m^3 por parcela) de uma espécie florestal no final da rotação economicamente aconselhável — a qual foi submetida a quatro diferentes tipos de manejo M1, M2, M3 e M4 —, efetuar o teste de Tukey para o nível de significância $\alpha = 0,05$.

Tabela 3.2.4. Volume de madeira em m³ por parcela (dados simulados).

Repetições	Tratamentos			
	M1	M2	M3	M4
1	34,5	45,0	43,5	42,0
2	38,3	53,3	54,0	36,8
3	43,5	54,0	51,8	44,3
4	41,3	51,0	50,3	42,8
5	39,5	49,6	53,4	---
6	44,0	52,3	---	---
Total	241,1	305,2	253,0	165,9
Médias	40,18333	50,8667	50,6000	41,4750

Tabela 3.2.5. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	3	532,7803	177,5934	13,71**
Resíduo	17	220,1692	12,9511	
Total	23	752,9495		

Teste de comparação múltipla de Tukey:

$$[(\hat{m}_i - \hat{m}_u) \pm \frac{q}{\sqrt{2}} \sqrt{QM_{RES}(\frac{1}{J_i} + \frac{1}{J_u})}]$$

Trat.	J_i	Médias
M2	6	50,8667
M3	5	50,6000
M4	4	41,4750
M1	6	40,1833

a) Comparando as médias dos tratamentos (M2 e M3):

$$\Delta = q_{(4; 17; 0,05)} \sqrt{\frac{1}{2} QM_{RES}(\frac{1}{6} + \frac{1}{5})} = 4,02 \sqrt{\frac{1}{2} (12,9511)(\frac{1}{6} + \frac{1}{5})} = 6,1944$$

Dado que $\hat{m}_2 - \hat{m}_3 = 0,2667 < 6,1944$, conclui-se, ao nível de significância $\alpha = 0,05$, pela não rejeição de $H_0 : m_2 = m_3$.

b) Comparando as médias dos tratamentos (M2 e M4):

$$\Delta = q_{(4; 17; 0,05)} \sqrt{\frac{1}{2} QM_{RES}(\frac{1}{6} + \frac{1}{4})} = 4,02 \sqrt{\frac{1}{2} (12,9511)(\frac{1}{6} + \frac{1}{4})} = 6,6033$$

Dado que $\hat{m}_2 - \hat{m}_4 = 9,3917 > 6,6033$, conclui-se, ao nível de significância $\alpha = 0,05$, pela rejeição de $H_0 : m_2 = m_4$.

c) Comparando as médias dos tratamentos M2 e M1:

$$\Delta = q_{(4; 17; 0,05)} \sqrt{\frac{QM_{RES}}{6}} = 4,02 \sqrt{\frac{12,9511}{6}} = 5,9061$$

Dado que $\hat{m}_2 - \hat{m}_1 = 10,6834 > 5,9061$, conclui-se, ao nível de significância $\alpha = 0,05$, pela rejeição de $H_0 : m_2 = m_1$.

d) Comparando as médias dos tratamentos (M3 e M4):

$$\Delta = q_{(4; 17; 0,05)} \sqrt{\frac{1}{2} QM_{RES} \left(\frac{1}{5} + \frac{1}{4}\right)} = 4,02 \sqrt{\frac{1}{2} (12,9511) \left(\frac{1}{5} + \frac{1}{4}\right)} = 6,8623$$

Dado que $\hat{m}_3 - \hat{m}_4 = 9,1250 > 6,8623$, conclui-se, ao nível de significância $\alpha = 0,05$, pela rejeição de $H_0 : m_3 = m_4$.

e) Comparando as médias dos tratamentos (M3 e M1):

$$\Delta = q_{(4; 17; 0,05)} \sqrt{\frac{1}{2} QM_{RES} \left(\frac{1}{6} + \frac{1}{5}\right)} = 4,02 \sqrt{\frac{1}{2} (12,9511) \left(\frac{1}{6} + \frac{1}{5}\right)} = 6,1944$$

Dado que $\hat{m}_3 - \hat{m}_1 = 10,4167 > 6,1944$, conclui-se, ao nível de significância $\alpha = 0,05$, pela rejeição de $H_0 : m_3 = m_1$.

f) Comparando as médias dos tratamentos (M4 e M1):

$$\Delta = q_{(4; 17; 0,05)} \sqrt{\frac{1}{2} QM_{RES} \left(\frac{1}{6} + \frac{1}{4}\right)} = 4,02 \sqrt{\frac{1}{2} (12,9511) \left(\frac{1}{6} + \frac{1}{4}\right)} = 6,6033$$

Dado que $\hat{m}_4 - \hat{m}_1 = 1,2917 < 6,6033$, conclui-se, ao nível de significância $\alpha = 0,05$, pela não rejeição de $H_0 : m_4 = m_1$.

Trat.	j_i	Médias	
M2	6	50,8667	a
M3	5	50,6000	a
M4	4	41,4750	b
M1	6	40,1833	b

Médias seguidas pela mesma letra, com o nível de significância $\alpha = 0,05$, não diferem pelo teste de Tukey.

A tabela 3.2.6 apresenta a análise de variância realizada pelo programa SAS, **Prog11**, em anexo no apêndice B:

Tabela 3.2.6. ANOVA realizada pelo SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Trat.	3	532,7803571	177,5934524	13,71	< 0,0001
Resíduo	17	220,1691667	12,9511275		
Total	20	752,9495238			

Alfa (α)	0,05
Graus de liberdade do resíduo	17
Quadrado médio do resíduo	12,95113
Valor crítico da amplitude total estudentizada	4,01999

Comparações entre Trat.	Diferenças entre médias	Intervalo de Confiança (1- $\hat{\alpha}$)		
M2 - M3	0,267	-5,928	6,461	
M2 - M4	9,392	2,788	15,995	*
M2 - M1	10,683	4,777	16,589	*
M3 - M4	9,125	2,263	15,987	*
M3 - M1	10,417	4,222	16,611	*
M4 - M1	1,292	-5,312	7,895	

3.3. MÉTODO DE SCHEFFÉ

Dado um experimento composto de I tratamentos, em que seja considerada a hipótese:

$$H_0 : Y_h = 0 \text{ vs } H_1 : Y_h \neq 0$$

Scheffé (1953) demonstra que os intervalos de confiança para todos os contrastes da forma (Y_h) , considerando um coeficiente de confiança $(1-\alpha)$, são dados por:

$$IC : [\hat{Y}_h \pm S]$$

$$Y_h = \sum_{i=1}^I c_i m_i \quad \left(\sum_{i=1}^I c_i = 0 \right)$$

Em que: $S = \sqrt{(I-1) \times F \times \hat{V}(\hat{Y}_h)}$

F : é o valor ao nível de significância α da distribuição "F" de Snedecor, correspondente aos graus de liberdade de tratamentos e para os graus de liberdade do resíduo.

- i) Para todos os possíveis intervalos, o coeficiente de confiança conjunto é $1-\alpha$;
- ii) O método de Scheffé é exato mesmo para experimentos não balanceados.

Pimentel Gomes (2000) menciona que o teste de Scheffé só deve ser aplicado quando o teste F , na comparação dos tratamentos, tenha sido significativo. No caso de apresentar resultado não significativo, nenhum contraste será significativo. O autor considera desaconselhável aplicar o teste para efetuar comparações entre duas médias, mas considera conveniente para testar contrastes complicados.

Exercício 3.3.1. Seja considerado um ensaio inteiramente ao acaso balanceado, referente à produção de madeira (m^3 por parcela) de uma espécie florestal, no final da rotação economicamente aconselhável, submetida a quatro diferentes tipos de manejo: M1, M2, M3 e M4.

Tabela 3.3.1. Volume em m^3 por parcela (dados simulados).

Trat.	Repetições					
	1	2	3	4	5	6
M1	34,5	38,3	43,5	41,3	39,5	44,0
M2	45,0	53,3	54,0	51,0	49,6	52,3
M3	43,5	54,0	51,8	50,3	53,4	45,6
M4	42,0	36,8	44,3	42,8	43,0	34,1

- a) Aplicar o teste de Scheffé para os contrastes entre duas médias (m_i e m_u):

Tabela 3.3.2. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	3	600,9346	200,3115	13,72**
Resíduo	20	292,0150	14,6007	
Total	23	892,9496		

$$\hat{V}(\hat{Y}_{iu}) = \left(\frac{c_1^2}{J_i} + \frac{c_2^2}{J_u} \right) \sigma^2 \quad \hat{V}(\hat{Y}_{iu}) = \left(\frac{1^2}{6} + \frac{1^2}{6} \right) \times 14,6007 = 4,8669$$

$$S_{iu} = \sqrt{(I-1) \times F \times \hat{V}(\hat{Y}_{iu})} = \sqrt{(4-1) \times 3,0984 \times 4,8669} = 6,7260$$

$$F_{3; 20; 0,05} = 3,0984 \quad F_{3; 20; 0,01} = 4,94$$

$\hat{m}_2 = 50,8667$	$\hat{m}_3 = 49,7667$	$\hat{m}_4 = 40,5000$	$\hat{m}_1 = 40,1833$
-----------------------	-----------------------	-----------------------	-----------------------

Contrastes		
M2 – M3	1,1000	ns
M2 – M4	10,3667	**
M2 – M1	10,6834	**
M3 – M4	9,2667	**
M3 – M1	9,5834	**
M4 – M1	0,3167	ns

Análise realizada pelo programa SAS, **Prog12**, em anexo no apêndice B.

Alfa (α)	0,05
Graus de liberdade do resíduo	20
Quadrado médio do resíduo	14,60075
Valor crítico de F	3,09839
Diferença mínima de significância	6,726

Scheffé (Agrupamento)	Médias	N	Trat.
A	50,867	6	M2
A	49,767	6	M3
B	40,500	6	M4
B	40,183	6	M1

b) Aplicar o teste de Scheffé, considerando que foram perdidas três parcelas, para os contrastes:
 $\hat{Y}_1 = \hat{m}_1 + \hat{m}_2 + \hat{m}_3 - 3\hat{m}_4$ e $\hat{Y}_2 = \hat{m}_2 - \hat{m}_3$.

Tabela 3.3.3. Volume em m³ por parcela.

Trat.	Repetições						Médias
	1	2	3	4	5	6	
M1	34,5	38,3	43,5	41,3	39,5	44,0	40,1833
M2	45,0	53,3	54,0	51,0	-	-	50,8250
M3	43,5	54,0	51,8	50,3	53,4	-	50,6000
M4	42,0	36,8	44,3	42,8	43,0	34,1	40,5000

Tabela 3.3.4. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Tratamentos	3	552,2165	184,0722	11,96**	0,0002
Resíduo	17	267,5158	15,7362		
Total	20	819,7324			

Tabela 3.3.5. ANOVA no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Trat.	3	552,2165476	184,0721825	11,70	0,0002
Resíduo	17	267,5158333	15,7362255		
Total	20	819,7323810			

b1) Comparando o grupo de manejo M1, M2, M3 com o M4:

$$\hat{Y}_1 = \hat{m}_1 + \hat{m}_2 + \hat{m}_3 - 3\hat{m}_4$$

$$\hat{Y}_1 = 40,1833 + 50,8250 + 50,60 - 3(40,50) = -20,1083 \text{ m}^3$$

$$\hat{V}(\hat{Y}_1) = \left(\frac{c_1^2}{J_1} + \frac{c_2^2}{J_2} + \frac{c_3^2}{J_3} + \frac{c_4^2}{J_4} \right) \sigma^2$$

$$\hat{V}(\hat{Y}_1) = \left[\frac{1^2}{6} + \frac{1^2}{4} + \frac{1^2}{5} + \frac{(-3)^2}{6} \right] \times 15,7362 = 33,3083$$

$$S_1 = \sqrt{(I-1) \times F \times \hat{V}(\hat{Y}_1)} = \sqrt{(4-1) \times 3,1968 \times 33,3083} = 17,8729$$

$$F_{3;17;0,05} = 3,1968$$

Como $|\hat{Y}_1| = |-20,1083| > S_1 = 17,8729$, conclui-se que o contraste é significativo, indicando que deve ser rejeitada a hipótese, $H_0: m_1 + m_2 + m_3 - 3m_4 = 0$, com um nível de significância $\alpha = 0,05$.

b2) Comparando os planos de manejo M2 e M3:

$$\hat{Y}_2 = \hat{m}_2 - \hat{m}_3 \quad \hat{Y}_2 = 50,8250 - 50,60 = 0,2250$$

$$\hat{V}(\hat{Y}_2) = \left(\frac{c_1^2}{J_1} + \frac{c_2^2}{J_2} \right) \sigma^2 \quad \hat{V}(\hat{Y}_2) = \left(\frac{1^2}{4} + \frac{1^2}{5} \right) \times 15,7362 = 7,0813$$

$$S_2 = \sqrt{(I-1) \times F \times \hat{V}(\hat{Y}_2)} = \sqrt{(4-1) \times 3,1968 \times 7,0813} = 8,2409$$

$$F_{3;17;0,05} = 3,1968.$$

Como $|\hat{Y}_2| = |0,2250| < S_2 = 8,2409$, conclui-se que o contraste é não significativo, indicando não rejeitar a hipótese, $H_0: m_2 = m_3$, com um nível de significância $\alpha = 0,05$.

3.4. MÉTODO DE DUNCAN

O método de Duncan (1955) consiste em ordenar as médias em ordem decrescente ou crescente, que possuam o mesmo número de repetições, abrangendo contrastes representados pela caracterização de uma posição no rol, a qual pode assumir os valores $2, 3, \dots, I$. Nesse caso, o nível de significância α varia de acordo com o posicionamento das duas médias na construção do contraste. O método foi construído para testar qualquer contraste entre duas médias de tratamentos, tendo-se:

$$H_0 : m_i - m_u = 0 \quad H_1 : m_i - m_u \neq 0 \quad i \neq u$$

O nível de significância α' foi desenvolvido por Duncan, visando obter um nível de proteção para α em todas as comparações entre as médias ordenadas, tal que α' é igual a $(1 - \alpha)^{(I^* - 1)}$. Quer dizer, varia de acordo com o número de médias ordenadas (I^*), abrangendo o contraste.

i) Trata-se de uma técnica sequencial, baseada na amplitude total estudentizada, válida para todos os contrastes entre duas médias;

ii) O método requer o balanceamento dos dados;

iii) O método considera que as médias sejam independentes, quando na verdade estão ordenadas.

De acordo com o método, rejeita-se H_0 para um nível de significância α' , se:

$$|\hat{m}_i - \hat{m}_u| \geq Z_{(I^*; n'; \alpha)} \sqrt{\frac{QMRes}{J}}$$

$Z_{I^*; n'; \alpha}$: é a amplitude total estudentizada, em que I^* é o número de médias ordenadas abrangidas pelo contraste e n' é o número de graus de liberdade do resíduo.

Pimentel Gomes (2000) cita que o teste de Duncan é um meio termo entre o rigor excessivo do teste de Tukey e a falta de rigor exagerada do teste “ t ”.

Exercício 3.4.1. Em um ensaio referente à produção de madeira (m^3 por parcela) de uma espécie florestal, no final da rotação economicamente aconselhável, submetida a quatro diferentes tipos de manejo (M1, M2, M3 e M4), efetuar o teste de Duncan.

Tabela 3.4.1. Volume em m^3 por parcela (dados simulados).

Trat.	Repetições					
	1	2	3	4	5	6
M1	34,5	38,3	43,5	41,3	39,5	44,0
M2	45,0	53,3	54,0	51,0	49,6	52,3
M3	43,5	54,0	51,8	50,3	53,4	45,6
M4	42,0	36,8	44,3	42,8	43,0	34,1

A tabela 3.4.2 apresenta a análise de variância para os quatro tipos de manejo.

Tabela 3.4.2. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	3	600,9346	200,3115	13,72**
Resíduo	20	292,0150	14,6007	
Total	23	892,9496		

$$F_{3; 20; 0,05} = 3,10 \quad F_{3; 20; 0,01} = 4,94$$

$\hat{m}_2 = 50,8667$	$\hat{m}_3 = 49,7667$	$\hat{m}_4 = 40,5000$	$\hat{m}_1 = 40,1833$
-----------------------	-----------------------	-----------------------	-----------------------

a) Obtenção das faixas críticas (Critical range) para as médias:

a1) Para contrastes que abrangem duas médias ($l^* = 2$):

$$D_2 = 2,95 \times \sqrt{\frac{14,6007}{6}} = 4,6018 \quad z_{2; 20; 0,05} = 2,95$$

a2) Para contrastes que abrangem três médias ($l^* = 3$):

$$D_3 = z_3 \times s(\hat{m}) = 3,097 \times \frac{\sqrt{14,6007}}{\sqrt{6}} = 4,8311 \quad z_{3; 20; 0,05} = 3,097$$

a3) Para contrastes que abrangem quatro médias ($l^* = 4$):

$$D_4 = z_4 \times s(\hat{m}) = 3,19 \times \frac{\sqrt{14,6007}}{\sqrt{6}} = 4,9762$$

b) Nas comparações, considerando as médias dos tratamentos na ordem decrescente, tal que (*) simboliza significância, em que $\alpha = 0,05$.

$\hat{m}_2 = 50,8667$	$\hat{m}_3 = 49,7667$	$\hat{m}_4 = 40,5000$	$\hat{m}_1 = 40,1833$
-----------------------	-----------------------	-----------------------	-----------------------

$\hat{Y}_{23} = \hat{m}_2 - \hat{m}_3 = 50,8667 - 49,7667 = 1,1000 \text{ ns}$	$\hat{Y}_{23} < D_2 = 4,6018$
$\hat{Y}_{24} = \hat{m}_2 - \hat{m}_4 = 50,8667 - 40,5000 = 10,3667 *$	$\hat{Y}_{24} > D_3 = 4,8311$
$\hat{Y}_{21} = \hat{m}_2 - \hat{m}_1 = 50,8667 - 40,1833 = 10,6834 *$	$\hat{Y}_{21} > D_4 = 4,9762$
$\hat{Y}_{34} = \hat{m}_3 - \hat{m}_4 = 49,7667 - 40,5000 = 9,2667 *$	$\hat{Y}_{34} > D_2 = 4,6018$
$\hat{Y}_{31} = \hat{m}_3 - \hat{m}_1 = 49,7667 - 40,1833 = 9,5834 *$	$\hat{Y}_{31} > D_3 = 4,8311$
$\hat{Y}_{41} = \hat{m}_4 - \hat{m}_1 = 40,5000 - 40,1833 = 0,3167 \text{ ns}$	$\hat{Y}_{41} < D_2 = 4,6018$

Trat.	Médias	
M2	50,8667	a
M3	49,7667	a
M4	40,5000	b
M1	40,1833	b

Médias seguidas pela mesma letra, ao nível de significância $\alpha = 0,05$, não diferem pelo teste de Duncan.

No teste de Duncan, cada contraste testado tem uma probabilidade $(1 - \alpha)^{I^*-1}$ de “não exibir como significativo uma diferença que na verdade é nula”, em que α é o nível de significância e I^* é o número de médias abrangidas no contraste. Considerando $\alpha = 0,05$ e $I^* = 2, 3, 4$, tem-se:

$$I^* = 2 \rightarrow (1 - \alpha)^{I^*-1} = (1 - 0,05)^1 = 0,95 \text{ (95 \%)}$$

$$I^* = 3 \rightarrow (1 - \alpha)^{I^*-1} = (1 - 0,05)^2 = 0,9025 \text{ (90,25 \%)}$$

$$I^* = 4 \rightarrow (1 - \alpha)^{I^*-1} = (1 - 0,05)^3 = 0,8574 \text{ (85,74 \%)}$$

É importante observar que o valor da diferença mínima de significância (Δ) de Tukey é maior que qualquer valor dos D_i de Duncan. Considerando que as análises efetuadas correspondem ao mesmo experimento, ocorre, para o teste de Tukey $\Delta = 6,1748$, enquanto para o teste de Duncan: $D_2 = 4,6018$, $D_3 = 4,8311$ e $D_4 = 4,9762$.

Os resultados apresentados nos exercícios 3.2.1 e 3.4.1 confirmam que o teste de Tukey é mais rigoroso que o de Duncan.

Análise realizada pelo programa SAS, **Prog13**, em anexo no apêndice B:

Duncan (Agrupamento)	Médias	N	Trat.
A	50,867	6	M2
A	49,767	6	M3
B	40,500	6	M4
B	40,183	6	M1

Número de médias	2	3	4
Faixas críticas	4,602	4,830	4,976

3.5. MÉTODO SNK (STUDENT, NEWMAN, KEULS)

O teste denominado SNK é também utilizado para testar os contrastes entre duas médias, aplicando a mesma metodologia de Duncan. Entretanto, usa a distribuição “ q ” do método Tukey, denominado amplitude total estudentizada.

$$[(\hat{m}_i - \hat{m}_u) \pm q_{(I, n-I)} \sqrt{\frac{QM_{RES}}{J}}] \quad H_0 : m_i = m_u \text{ vs } H_1 : m_i \neq m_u$$

$$(n = I \times J)$$

3.6. MÉTODO DE DUNNETT

Este método foi construído com o objetivo de comparar os tratamentos ministrados com um tratamento controle ou padrão (testemunha):

$$H_0 : m_i = m_0 \text{ VS } H_1 : m_i \neq m_0$$

Dunnett (1955) apresentou o seguinte intervalo de confiança para a totalidade dos contrastes:

$$\Pr[\hat{Y} \pm t_{(I-1; n-I; \alpha)} \sqrt{\hat{V}(\hat{Y})}] = 1 - \alpha$$

$$\hat{V}(\hat{m}_0 - \hat{m}_i) = 2 \frac{QM_{RES}}{J}$$

$t_{I-1; I-n; \alpha}$: valor tabelar da distribuição de Dunnett;

$I-1$: número de tratamentos, excetuando a testemunha;

n : número total de observações;

$I-n$: número de graus de liberdade do resíduo.

i) O método é exato quando os dados são balanceados;

ii) Ele é mais sensível que os métodos de Tukey e Scheffé.

Exercício 3.6.1. Em um ensaio, referente à produção de madeira (m³ por parcela) de uma espécie florestal que foi submetida a quatro diferentes tipos de manejo — M1(Testemunha), M2, M3 e M4 —, o delineamento experimental adotado foi inteiramente casualizado. Considerando que os dados (simulados) obtidos referentes ao volume de madeira, em m³ por parcela, no final da rotação economicamente aconselhável, estão apresentados na tabela 3.6.1, efetuar o teste de Dunnett.

Tabela 3.6.1. Volume (m³ por parcela) no final da rotação.

Trat.	Repetições					
	1	2	3	4	5	6
M1(T)	34,5	38,3	43,5	41,3	39,5	44,0
M2	45,0	53,3	54,0	51,0	49,6	52,3
M3	43,5	54,0	51,8	50,3	53,4	45,6
M4	42,0	36,8	44,3	42,8	43,0	34,1

Tabela 3.6.2. ANOVA.

F. V.	G.L.	S.Q.	Q.M.	F
Tratamentos	3	600,9346	200,3115	13,72**
Resíduo	20	292,0150	14,6007	
Total	23	892,9495		

a) Estimativa da variância do contraste entre duas médias:

$$\hat{V}(\hat{Y}) = \left(\frac{1}{J_o} + \frac{1}{J_i}\right) QM_{RES} = \left(\frac{1}{6} + \frac{1}{6}\right) \times 14,6007 = 4,8669$$

b) Estimativa do erro padrão do contraste:

$$s(\hat{Y}) = \sqrt{\hat{V}(\hat{Y})} = \sqrt{4,8669} = 2,2061$$

c) Critério do teste:

$$\Pr[\hat{Y} \pm t_{(I-1; n-I; \alpha)} \sqrt{\hat{V}(\hat{Y})}] = 1 - \alpha$$

$$t_{3; 20; 0,05} = 2,54043 \quad n - I = 24 - 4 = 20$$

$$\Pr[\hat{Y} \pm t_{(3; 20; 0,05)} \sqrt{\hat{V}(\hat{Y})}] = \Pr(\hat{Y} \pm 2,54043 \times 2,2061) = 0,95$$

$$\Pr(\hat{Y} \pm 5,6045) = 0,95$$

$\hat{m}_2 = 50,8667$	$\hat{m}_3 = 49,7667$	$\hat{m}_4 = 40,5000$	$\hat{m}_1 = \hat{m}_0 = 40,1833$
-----------------------	-----------------------	-----------------------	-----------------------------------

$\hat{m}_2 - \hat{m}_0 = 10,6834 > 5,6045$	*
$\hat{m}_3 - \hat{m}_0 = 9,5834 > 5,6045$	*
$\hat{m}_4 - \hat{m}_0 = 0,3167 < 5,6045$	ns

Análise realizada pelo programa SAS, **Prog14**, em anexo no apêndice B:

Alfa (α)	0,05
Graus de liberdade do resíduo	20
Quadrado médio do resíduo	14,60075
Valor crítico de Dunnett's "t"	2,54043
Diferença mínima de significância	5,6045

Comparações entre médias de Trat.	Diferenças entre médias	Intervalo de Confiança (1 - α)		
M2 - M1	10,683	5,079	16,288	*
M3 - M1	9,583	3,979	15,188	*
M4 - M1	0,317	-5,288	5,921	

3.7. MÉTODO DE BONFERRONI

A correção de Bonferroni define um novo valor de significância que leva em consideração o número de testes realizados, o qual determina mais rigor em aceitar a rejeição da hipótese de nulidade. O método é indicado para um determinado número fixo de contrastes e pode ser utilizado para analisar experimentos não balanceados.

$$H_0 : Y_h = 0 \text{ vs } H_1 : Y_h \neq 0$$

Seja considerado p o número de comparações de médias, duas a duas, então, se houver I tratamentos, podem existir $p = \binom{I}{2} = \frac{I(I-1)}{2}$ comparações. O teste de Bonferroni utiliza $\frac{\alpha}{p}$ como nível de significância para um erro experimental α . O intervalo de confiança para o contraste Y_h é definido como:

$$[\hat{Y}_h \pm (t_{n-I; \frac{\alpha}{p}}) \times \sqrt{\hat{V}(\hat{Y}_h)}] \quad h = 1, \dots, p$$

p : número de contrastes;

$t_{\frac{\alpha}{p}; n-I}$: é o quantil de ordem $(1 - \frac{\alpha}{2p})$ da distribuição “ t ” de Student com $n - I$ graus de liberdade.

i) A probabilidade de que os p intervalos estejam corretos é maior ou igual a $1 - \alpha$;

ii) O método é mais preciso quanto menor for o número de contrastes testados.

Alguns métodos de comparação múltipla não conseguem manter, no valor especificado α , o erro tipo I global constante, que é o valor original para rejeitar a hipótese nula, definido normalmente como 0,05.

O método de Bonferroni propõe uma forma de correção chamada correção de Bonferroni, isto é, que define um novo valor de significância em função do número p de comparações realizadas. Quer dizer, $\alpha' = \frac{\alpha}{p}$, resultando em $\alpha = p\alpha'$ para o conjunto.

Por exemplo, se em quatro tratamentos estão sendo comparadas as médias, duas a duas, e se o máximo erro tipo global permitido é $\alpha = 0,05$, resultarão em seis pares de médias ($p = \frac{4 \times 3}{1 \times 2} = 6$) para os testes. Consequentemente a correção de Bonferroni apresentará o seguinte valor para $\alpha' = \frac{0,05}{6} = 0,008333$.

Considerando que o teste “ t ” só deve ser aplicado quando os contrastes forem ortogonais e previamente escolhidos, ou seja, antes de serem observados os dados — assim como em número no máximo igual ao grau de liberdade de tratamentos —, o uso do teste de Bonferroni pode ser uma alternativa.

Pimentel Gomes (2000) comenta que o teste de Bonferroni é um aperfeiçoamento do teste “ t ”. O teste “ t ”, mesmo com as restrições já citadas, não é exato. Para contornar esse problema, Bonferroni indicou o uso, para cada contraste, de um nível de probabilidade $\alpha' = \frac{\alpha}{p}$. Por outro lado, o autor considera ser o teste de Bonferroni adequado para aplicar em contrastes complexos, bem como para pequeno número de contrastes, mas excessivamente conservador quando esse número cresce.

A correção de Bonferroni tende a inflar o erro do tipo II, principalmente quando se compara um número muito grande de testes. Assim, tal erro evita a identificação de diferenças que possam existir.

A correção de Bonferroni é usada para diminuir o erro tipo I (rejeitar H_0 quando H_0 é verdadeira). Essa correção reduz a chance de concluir que há diferença estatística significativa, quando na realidade não houver. Por outro lado, na análise de um experimento, é importante refletir acerca da diminuição do erro tipo I, o qual corresponde a aumentar a chance do erro tipo II (aceitar H_0 quando H_0 é falsa), significando que o resultado não é estatisticamente significativo quando é. Ou seja, ressalta a importância de se manter o equilíbrio entre os erros tipo I e II.

Exercício 3.7.1. Em um ensaio referente à produção de madeira (m^3 por parcela) de uma espécie florestal no final da rotação economicamente aconselhável, submetida a quatro diferentes tipos de manejo (M1, M2, M3 e M4), o delineamento experimental adotado foi inteiramente casualizado, em que os dados obtidos, tabela 3.7.1, são referentes ao volume de madeira no final da rotação. Nesse contexto, aplicar o teste de Bonferroni.

$$[\hat{Y}_j \pm (t_{n-I, \frac{\alpha}{p}}) \times \sqrt{\hat{V}(\hat{Y}_j)}]$$

$$p = \frac{I(I-1)}{2} = \frac{4(4-1)}{2} = 6 \rightarrow \frac{\alpha}{6} = 0,0083333 \rightarrow (1 - \frac{\alpha}{p}) = 0,9916667$$

Tabela 3.7.1. Volume m^3 por parcela (dados simulados).

Trat.	Repetições					
	1	2	3	4	5	6
M1	34,5	38,3	43,5	41,3	39,5	44,0
M2	45,0	53,3	54,0	51,0	49,6	52,3
M3	43,5	54,0	51,8	50,3	53,4	45,6
M4	42,0	36,8	44,3	42,8	43,0	34,1

Tabela 3.7.2. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	3	600,9346	200,3115	13,72**
Resíduo	20	292,0150	14,6007	
Total	23	892,9496		

Da tabela A12, considerando seis comparações: $t_{20; 0,008333} = 2,92712$.

$$\hat{V}(\hat{Y}) = \frac{2}{J} QM_{RES} = \left(\frac{2 \times 14,60075}{6} \right) = 4,8669$$

$$[\hat{Y}_j \pm (t_{n-I; \frac{\alpha}{2p}}) \times \sqrt{\hat{V}(\hat{Y}_j)}] = (Y_j \pm 2,92712 \times \sqrt{4,8669})$$

$$(\hat{Y}_j \pm 6,4575) = 0,95$$

$\hat{m}_2 = 50,8667$	$\hat{m}_3 = 49,7667$	$\hat{m}_4 = 40,5000$	$\hat{m}_1 = 40,1833$
-----------------------	-----------------------	-----------------------	-----------------------

$\hat{m}_2 - \hat{m}_3 = 1,1000 < 6,4575$	<i>ns</i>
$\hat{m}_2 - \hat{m}_4 = 10,3667 > 6,4575$	*
$\hat{m}_2 - \hat{m}_1 = 10,6834 > 6,4575$	*
$\hat{m}_3 - \hat{m}_4 = 9,2667 > 6,4575$	*
$\hat{m}_3 - \hat{m}_1 = 9,5834 > 6,4575$	*
$\hat{m}_4 - \hat{m}_1 = 0,3167 < 6,4575$	<i>ns</i>

Análises realizadas pelo programa SAS, **Prog14**, em anexo no apêndice B:

Alfa (α)	0,05
Graus de liberdade do resíduo	20
Quadrado médio do resíduo	14,60075
Valor crítico de “t” de Bonferroni	2,92712
Diferença mínima de significância	6,4575

Bonferroni (Agrupamento)	Médias	N	Trat.
A	50,867	6	M2
A	49,767	6	M3
B	40,500	6	M4
B	40,183	6	M1

3.8. COMPARAÇÕES ENTRE OS TESTES DE TUKEY, SCHEFFÉ, BONFERRONI E DUNCAN

Para verificar o rigor dos testes de Tukey, Scheffé, Bonferroni e Duncan — sejam seus intervalos de confiança, considerando o ensaio referente à produção de madeira (m^3 por parcela) da espécie florestal que foi submetida a quatro diferentes tipos de manejo (M1, M2, M3 e M4) —, no caso de o intervalo conter o valor zero (0), deve-se aceitar H_0 . Em caso contrário, rejeite-se H_0 .

Sejam as médias referentes aos quatro métodos de manejo:

$\hat{m}_2 = 50,8667$	$\hat{m}_3 = 49,7667$	$\hat{m}_4 = 40,5000$	$\hat{m}_1 = 40,1833$
-----------------------	-----------------------	-----------------------	-----------------------

Tabela 3.8.1. Comparações entre os testes de Tukey, Scheffé, Bonferroni e Duncan.

Contrastes	$\hat{Y}_i$	Intervalos de confiança			
		Tukey	Scheffé	Bonferroni	Duncan
M2 – M3	1,1	(-5,074; 7,275)	(-5,626; 7,826)	(-5,357; 7,557)	(-3,502; 5,702)
M2 – M4	10,367	(4,192; 16,541)	(3,641; 17,093)	(3,909; 16,824)	(5,537; 15,197)
M2 – M1	10,683	(4,5096; 16,858)	(3,957; 17,409)	(4,226; 17,141)	(5,707; 15,659)
M3 – M4	9,267	(3,092; 15,441)	(2,541; 15,993)	(2,809; 15,724)	(4,665; 13,869)
M3 – M1	9,583	(3,409; 15,758)	(2,857; 16,309)	(3,126; 16,041)	(4,753; 14,413)
M4 – M1	0,317	(-5,858; 6,491)	(-6,409; 7,043)	(-6,141; 6,774)	(-4,285; 4,919)
		$\Delta = 6,1748$	$\Delta = 6,726$	$\Delta = 6,4575$	

Considerando as amplitudes dos intervalos de confiança (I.C.) para os testes:

Tukey	Scheffé	Bonferroni	Duncan
12,3496	13,4520	12,9150	$[I^* = 2 (9,2036), I^* = 3 (9,6622), I^* = 4 (9,9524)]$

Duncan ($D_2 = 4,6018$, $D_3 = 4,8311$ e $D_4 = 4,9762$).

Os resultados mostram que, em comparação com os demais métodos, o teste de Duncan tem uma amplitude menor do intervalo de confiança. O teste de Tukey, por sua vez, tem uma amplitude menor que os de Scheffé e Bonferroni. Já o teste de Bonferroni apresenta uma amplitude menor que o teste de Scheffé.

3.9. CÁLCULO DO NÚMERO DE REPETIÇÕES IDEAL

A determinação do número de repetições no planejamento de um experimento é um problema complexo. Existem várias soluções para estimar o número de repetições em um ensaio. Pimentel Gomes (2000) ressalta que nenhuma alternativa é inteiramente satisfatória, citando ainda que, na pesquisa em ciências agrárias, resultados razoáveis dificilmente são obtidos, por intermédio de ensaios que tenham menos de vinte parcelas. Ressalte-se também que, em geral, deve haver pelo menos dez graus de liberdade para o resíduo.

A partir de uma pesquisa bibliográfica de ensaios anteriores, implantados em condições análogas, pode obter-se uma estimativa prévia do desvio padrão ($\hat{\sigma}$), com o seu respectivo número de graus de liberdade (n_2), além de ser possível estabelecer uma diferença mínima significativa (d), a qual deverá ser estatisticamente comprovada pelo experimento.

Pimentel Gomes (2000) apresenta a seguinte fórmula, em que o número de repetições é obtido, fixando-se uma amplitude total estudentizada (q), sendo F , ao nível α de probabilidade escolhido, o valor correspondente com o número n_1 de graus de liberdade do novo experimento e n_2 referente aos graus de liberdade de ensaios anteriores.

$$J = \frac{q^2 \hat{\sigma}^2 F}{d^2} \quad F_{n_1; n_2; \alpha}$$

O número de repetições (J) possibilitará obter uma probabilidade α de que o experimento não venha a comprovar a diferença (d) estipulada, ou seja, uma probabilidade $(1-\alpha)$ que seja comprovada estatisticamente pelo teste de Tukey.

Uma vez que o número de repetições J está em função dos valores de q e F , obviamente, pode-se obter, a partir de uma tentativa inicial, uma solução por aproximações sucessivas.

Por outro lado, considerando o procedimento análogo, quando é conhecido o coeficiente de variação (CV) e a diferença mínima (d), em porcentagem, a fórmula apresenta seguinte formatação:

$$J = \frac{q^2 (CV)^2 F}{d^2}$$

No caso de um coeficiente de variação de 10 % em um experimento florestal anterior, o qual tenha sido obtido a partir de um quadrado médio de resíduo com 60 graus de liberdade, qual deverá ser o número de repetições de um novo ensaio, casualizado em blocos, com cinco tratamentos e uma diferença mínima $d = 20\%$?

a) Inicialmente, considerando quatro blocos, logo: $q_{5; 12; 0,05} = 4,51$ e $F_{12; 60; 0,05} = 1,92$.

F.V.	G.L.
Blocos	3
Tratamentos	4
Resíduo	12
Total	19

$$J = \frac{4,51^2 \times 10^2 \times 1,92}{20^2} = 9,76$$

O número de blocos recomendável estará entre o número testado (4) e o valor obtido (9,76). Assim sendo, o uso de quatro blocos é insuficiente, sugerindo a tentativa de seis blocos.

b) Considerando seis blocos, logo: $q_{5; 20; 0,05} = 4,23$ e $F_{20; 60; 0,05} = 1,75$.

F.V.	G.L.
Blocos	5
Tratamentos	4
Resíduo	20
Total	29

$$J = \frac{4,23^2 \times 10^2 \times 1,75}{20^2} = 7,8$$

O número de blocos recomendável deverá estar entre o número testado (6) e o valor obtido (7,8). Assim sendo, o uso de seis blocos é insuficiente, sugerindo, desta vez, testar sete blocos.

c) Considerando sete blocos, então $q_{5; 24; 0,05} = 4,17$ e $F_{24; 60; 0,05} = 1,70$.

F.V.	G.L.
Blocos	6
Tratamentos	4
Resíduo	24
Total	34

$$J = \frac{4,17^2 \times 10^2 \times 1,70}{20^2} = 7,4$$

O número de blocos recomendável estará entre o número testado (7) e o valor obtido (7,4). Assim sendo, o uso de sete blocos errará ligeiramente por falta. Por outro lado, o pesquisador utilizando oito blocos dimensionará por excesso. Outros métodos para a obtenção do número ideal de repetições estão citados em Pimentel Gomes (2000).

3.10. BIBLIOGRAFIA

BANZATTO, D. A.; KRONKA, N. **Experimentação agrícola**. Jaboticabal: FUNEP, 1989. 247 p.

DUNCAN, D. B. Multiple range and multiple F tests, **Biometrics**, v. 11, p. 1-42, 1955.

DUNNETT, C. W. A multiple comparison procedure for comparing several treatments with a control, **Journal of the American Statistical Association**, v.50, n.272, p. 1096 -1121, 1955.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 14ª ed. São Paulo: Nobel, 2000. 475 p.

SCHEFFÉ, H. A Method for judging all contrasts in Analysis of Variance. *Biometrika*, 40:87-104, 1953.

SILVA, I. P.; SILVA, J. A. A. **Métodos estatísticos aplicados à pesquisa científica: uma abordagem para profissionais da pesquisa agropecuária**. Universidade Federal Rural de Pernambuco: UFRPE, 2002. 324 p.

TUKEY, J. W. Comparing individual means in the analysis of variance, **Biometrics**, v. 5, p. 99-114, 1949b.

CAPÍTULO 4

DELINEAMENTO CASUALIZADO EM BLOCOS

Este capítulo apresenta definições, conceitos e aplicações sobre a análise de experimentos casualizados em blocos. Apresenta os quadros das análises de variância com os respectivos componentes para os modelos fixo, aleatório e misto. Discorre acerca da aplicação dos métodos de Kolmogorov-Smirnov, Cramér-Von Mises, Anderson-Darling e Shapiro-Wilk, para testar a normalidade dos resíduos do modelo linear, assim como o teste de Tukey para verificar a existência de aditividade dos efeitos do modelo. Apresenta o teste para comparar a efetividade do experimento casualizado em blocos com o inteiramente ao acaso. Também apresenta exemplo de aplicação de experimento casualizado em blocos com tratamentos comuns. Mostra o quadro da análise de variância para o ensaio casualizado em blocos com K amostras por parcela, apresentando os componentes de variâncias. Apresenta a análise de um experimento de adubação, considerando a regressão polinomial e a equação de Mitscherlich, para obter a dose economicamente aconselhável.

4.1. INTRODUÇÃO

O delineamento casualizado de blocos completos, também chamado de “blocos ao acaso” ou simplesmente “blocos casualizados”, foi proposto, em 1925, por Sir Ronald A. Fisher.

Por outro lado, a denominação “blocos casualizados” denota que os blocos são aleatórios, isto é, apresentam efeitos aleatórios, o que não é verdadeiro na maioria dos casos. O efeito aleatório dos blocos significa sortear um determinado número de blocos de uma população de blocos. Neste livro será adotada a designação “Delineamento Casualizado em Blocos (DCB)” ou “Delineamento Casualizado em Blocos Completos”, isto é, “Design randomized complete-block”.

O delineamento casualizado em blocos considera os princípios fundamentais da repetição, da casualização, assim como incorpora o princípio do controle local, que permite estratificar o ambiente em subambientes (blocos), possibilitando estabelecer condições homogêneas para os tratamentos dentro de cada bloco. O controle da variação ambiental pode promover a redução do erro experimental. Contudo, isso só ocorrerá se o ambiente experimental for relativamente heterogêneo, haja vista que o controle local reduz os graus de liberdade do erro e, caso a parte da variação explicada pelos blocos não seja considerável, a referida redução pode aumentar a estimativa do erro e prejudicar a precisão na comparação de médias de tratamentos.

Devido às suas características — como a facilidade de instalação e a caracterização de, geralmente, cada repetição ser um bloco —, o delineamento em blocos casualizados é o mais empregado nas pesquisas em ciências agrárias, pois permite a realização de experimentos no campo em áreas heterogêneas (separadas por diferenças na topografia e em tipos diferentes de solo), em que, dificilmente, é possível encontrar as condições ideais de homogeneidade ambiental, que são as requeridas no delineamento inteiramente ao acaso. Logo, os agrupamentos das unidades experimentais em subconjuntos mais homogêneos são denominados blocos.

Dentro de cada bloco (internamente) são distribuídos aleatoriamente todos os tratamentos envolvidos no experimento, resultando que a maior parte da variabilidade interna do conjunto de unidades experimentais será expressa pela heterogeneidade entre blocos. As unidades experimentais, dentro dos blocos, devem estar sob ambiente mais homogêneo possível (solo, topografia, irrigação etc.). Quanto maior a heterogeneidade entre os blocos maior será, em relação ao delineamento inteiramente ao acaso, a eficiência do delineamento casualizado em blocos.

A distribuição dos tratamentos nas parcelas é feita de forma casualizada, porém, com uma restrição: o sorteio dos tratamentos é feito apenas dentro de cada bloco. Ou seja, cada bloco deve conter todos os tratamentos (o que caracteriza os blocos completos). A vantagem mais destacada dos experimentos casualizados em blocos consiste em permitir o uso de unidades experimentais heterogêneas.

Ressalte-se que a denominação “Casualizado em blocos”, embora largamente utilizada, não é suficientemente precisa, visto que poderia incluir os delineamentos de blocos incompletos, tema que será tratado no capítulo 9.

Neste livro será considerado o delineamento casualizado em blocos com uma única repetição de cada tratamento dentro do bloco. Logo, o número de repetições deverá ser igual ao número de blocos.

Cada bloco é um subconjunto de parcelas homogêneas, no qual os tratamentos deverão manifestar seus efeitos de forma independente do bloco e de forma aditiva. Por exemplo, parcelas numa mesma altitude, com mesmo tipo de solo, de árvores de mesma procedência, altura, espécie, diâmetro etc. De um modo geral, os blocos podem diferir entre si em maior ou menor grau. Contudo, tais diferenças não podem causar interação entre os blocos e os tratamentos, pois essa interação irá inflacionar o erro experimental e reduzir a precisão dos resultados na análise do experimento. Em tese, se um tratamento se comportar melhor em um determinado bloco, deve ser esperado que também seja melhor nos demais blocos.

Nesse delineamento fica evidente a necessidade de aditividade entre tratamentos e blocos (os termos do modelo são somados). Qualquer falta de aditividade irá inflacionar os resíduos. É por essa razão é que o uso de blocos, quando necessário, faz-se tão importante. Na análise exploratória de dados é fácil detectar a falta de aditividade, simplesmente construindo um gráfico em linha, ou seja, plotando os blocos $B_1, B_2, \dots, B_j$ nos eixos das abscissas e as correspondentes respostas dos tratamentos (uma linha por bloco) nas ordenadas. Se ocorrer paralelismo das respostas, de bloco para bloco, haverá aditividade dos efeitos. Entretanto, se as linhas se cruzarem (interação entre efeitos de blocos e tratamentos), há fortes evidências da falta de aditividade. A aditividade pode ser avaliada pelo teste estatístico de Tukey (1949a).

O delineamento casualizado em blocos é o mais usado na pesquisa florestal de campo, visto que quase sempre há suspeita de alguma variação no ambiente experimental e o número de tratamentos, em geral, é baixo. Em experimentos de campo, um bloco consiste de parcelas homogêneas, lado a lado (contíguas), alinhadas ou não, com mesmo formato e área. Os blocos podem diferir bastante entre si, podendo, inclusive, estar separados espacialmente ou no tempo. Nesse delineamento, ou mesmo em outros que fazem uso de controle local, a instalação e/ou práticas culturais de condução do experimento, ou mesmo as observações (coleta de dados), devem ser feitas por bloco.

O uso do delineamento casualizado em blocos geralmente não é recomendável para testar um número elevado de tratamentos, haja vista que o tamanho do bloco aumenta proporcionalmente com o número de tratamentos, tornando difícil manter a homogeneidade das parcelas dentro dos blocos, consequentemente, comprometendo as comparações dos tratamentos. Nesses casos, em geral, espera-se a elevação do erro experimental com o aumento do número de tratamentos. Markus (1971) não recomenda utilizar a casualização nos blocos em experimentos com mais de 16 tratamentos. O dimensionamento do número de tratamentos varia de acordo com as condições ambientais em que o experimento será instalado e conduzido, bem como com o tamanho de parcela a ser empregado.

O delineamento casualizado em blocos requer um determinado controle da variação ambiental, reduzindo, assim, o erro experimental. Contudo, a redução do erro só ocorrerá se o ambiente experimental for relativamente heterogêneo, pois o controle local custa uma redução nos graus de liberdade do erro e, caso a parte da variação isolada para blocos não seja expressiva, tal redução pode elevar a estimativa do erro e prejudicar a precisão na comparação de médias de tratamentos.

Considerando que o delineamento casualizado em blocos seja mais eficiente do que o delineamento inteiramente casualizado, nem sempre os resultados da análise serão mais precisos. Logo, não se pode afirmar que o delineamento casualizado em blocos garanta experimentos mais precisos do que o delineamento inteiramente casualizado.

Em experimentos no campo, visando atingir o máximo de homogeneidade dentro dos blocos, é recomendável estruturar os blocos buscando a forma mais próxima possível de um quadrado. A desvantagem do delineamento casualizado em blocos é que a organização em blocos implica na redução do número de graus de liberdade para calcular a variância residual, assim como a exigência de homogeneidade dentro do bloco pode limitar o número de tratamentos. Por outro lado, se realmente ocorrer grande heterogeneidade entre os blocos, isso resultará no aumento significativo da soma de quadrados de blocos, o que acarretará diminuição do valor da soma de quadrados do resíduo e, conseqüentemente, de seu quadrado médio, aumentando, portanto, a eficiência da casualização dos tratamentos em blocos.

Padovani (2014) comenta que os blocos controlam uma causa de variação e estabelecem uma restrição à casualização. Essa restrição, devido à constituição dos blocos, segundo ele, indica para a não realização do teste estatístico para a causa de variação de blocos. Isso não faz sentido, pois trata de uma fonte da variação de controle e não do interesse para a comparação. Se a fonte, definida como bloco, tiver importância para estabelecer comparações na pesquisa, então o esquema de fatores se torna o procedimento mais adequado para a combinação de níveis dos dois fatores em estudo.

Uma desvantagem do delineamento casualizado em blocos é a condição de ter o mesmo número de repetições por tratamento. Essa exigência gera complicações adicionais à análise estatística, pois caso haja perda de parcelas, ocasionará desbalanceamento entre os tratamentos. Pimentel Gomes (2000) e Banzatto e Kronka (1989) mostram como obter as estimações das observações para a(s) parcela(s) perdida(s), possibilitando obter a solução analítica e os devidos ajustamentos. Por outro lado, no caso de várias parcelas perdidas, outra solução pode ser a eliminação de tratamento(s) ou bloco(s) que sofreram as perdas, de forma a garantir o balanceamento. Todavia, essa alternativa só é viável para experimentos com número alto de tratamentos e de repetições (situação rara, nas ciências agrárias, em experimentos casualizados em blocos).

Padovani (2014) resume e destaca, considerando o delineamento casualizado em blocos, os seguintes aspectos:

- a) A casualização ocorre dentro dos blocos (os blocos são estratos definidos quanto à heterogeneidade das unidades experimentais e, portanto, fixados como controles);
- b) Quanto maior a heterogeneidade entre blocos, maior deverá ser a eficiência do delineamento (a perda de heterogeneidade entre blocos indica a falta da necessidade de controle local);
- c) Os experimentos em blocos são feitos, essencialmente, a fim de comparar tratamentos (blocos não são construídos para teste estatístico, apenas por necessidade de controle de diferentes ambientes);
- d) É essencial que cada bloco forneça condição similar a cada tratamento (unidades dentro de blocos devem ser a única fonte de variação);

- e) Os blocos são completos quando cada bloco contém todos os tratamentos do estudo;
- f) Fazer blocos significa impor uma restrição da distribuição das unidades experimentais e a casualidade será restrita às unidades experimentais dentro de cada bloco;
- g) O autor cita exemplos de blocos: posição na estufa, solo, irrigação, idade, peso e lote de sementes;
- h) Todas as considerações estatísticas feitas para o delineamento inteiramente ao acaso são também válidas para os ensaios casualizados em blocos (aditividade dos efeitos, independência dos erros, variâncias homogêneas e normalidade dos dados).

No caso de experimentos instalados em área de campo, a implantação dos blocos deve considerar topografia, solo etc. Se o experimento for instalado em laboratório, por sua vez, o bloco poderá ser o operador, o aparelho medidor ou cada dia em que os resultados forem medidos. Há muitas outras situações em que blocos podem ser utilizados. É usual denominar a expressão “controle local” para definir um delineamento que seja instalado em blocos.

Considerar o exemplo de um planejamento, ilustrado na figura 4.1, com casualização em quatro blocos, para testar cinco procedências de uma espécie florestal (P1, P2, P3, P4, P5). As procedências deverão ser sorteadas aleatoriamente dentro de cada bloco (quatro sorteios diferentes):

Bloco 1	P3	P2	P5	P1	P4
Bloco 2	P4	P3	P1	P2	P5
Bloco 3	P1	P4	P5	P3	P2
Bloco 4	P5	P3	P2	P4	P1

Figura 4.1. Distribuição dos tratamentos, considerando o delineamento casualizado em blocos completos.

4.2. ANÁLISE DE VARIÂNCIA (ANOVA) E COMPONENTES DE VARIÂNCIAS

A análise de variância, no delineamento em blocos, requer o atendimento das condições clássicas de homoscedasticidade, aditividade, normalidade e independência para a ANOVA. A Tabela 4.2.1 apresenta as expressões para a obtenção da análise de variância.

A análise de variância, aqui denominada análise segundo dois critérios de classificação (tratamentos e blocos), obedece ao modelo a seguir, que pressupõe que a manifestação de uma dada variável resposta, no nível de cada parcela, é decorrente dos efeitos de causas controladas (tratamentos e blocos) e das causas não controladas no experimento (erro ou resíduo):

$$x_{ij} = m_i + b_j + e_{ij}, \text{ em que: } m_i = m + t_i \begin{cases} i=1, 2, \dots, I \\ j=1, 2, \dots, J \end{cases}$$

No caso do modelo com efeitos fixos:

$$E(x_{ij}) = m + t_i + b_j$$

$$V(x_{ij}) = \sigma^2$$

$$x_{ij} \sim N(m + t_i + b_j, \sigma^2)$$

Tabela 4.2.1. ANOVA para o delineamento casualizado em blocos completos.

F.V.	G.L.	S.Q.	Q.M.	F
Trat.	$I - 1$	$\frac{1}{J} \sum_{i=1}^I T_i^2 - C$	$\frac{SQ_{TRAT}}{I - 1}$	$\frac{QM_{TRAT}}{QM_{RES}}$
Blocos	$J - 1$	$\frac{1}{I} \sum_{j=1}^J B_j^2 - C$	$\frac{SQ_{BLOCOS}}{J - 1}$	$\frac{QM_{BLOCOS}}{QM_{RES}}$
Resíduo	$(I - 1)(J - 1)$	Diferença	$\frac{SQ_{RES}}{(I - 1)(J - 1)}$	
Total	$IJ - 1$	$\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C$		

No delineamento casualizado em blocos, a análise de variância consiste em desdobrar a soma de quadrados total (SQ_{TOTAL}), em três fontes de variação. A primeira referente aos tratamentos (SQ_{TRAT}); a segunda é relativa aos blocos (SQ_{BLOCOS}) e a terceira expressa as flutuações casuais (SQ_{RES}).

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C \quad SQ_{TRAT} = \frac{1}{J} \sum_{i=1}^I T_i^2 - C \quad SQ_{BLOCOS} = \frac{1}{I} \sum_{j=1}^J B_j^2 - C$$

$$C = \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ}$$

$$SQ_{RES} = SQ_{TOTAL} - (SQ_{TRAT} + SQ_{BLOCOS})$$

O teste de hipótese, relativo à tabela 4.2.1, pode ser obtido por meio de:

$$H_0 : \text{Não existem efeitos de tratamentos } H_0 : t_1 = t_2 = \dots = t_I \Leftrightarrow H_0 : m_1 = m_2 = \dots = m_I = m ;$$

$$H_1 : \text{Existem efeitos de tratamentos } H_1 : t_i \neq 0 \quad (i = 1, \dots, I).$$

Sob a veracidade de H_0 , a estatística $\frac{QM_{TRAT}}{QM_{RES}}$ tem distribuição F de Snedecor com $I - 1$ (graus de liberdade do numerador) e $(I - 1)(J - 1)$ (graus de liberdade do denominador). A regra de decisão é rejeitar H_0 se $F \geq F_{(I-1); (I-1)(J-1); \alpha}$. Não cabendo rejeição em caso contrário.

Seja o modelo matemático:

$$x_{ij} = m_i + b_j + e_{ij}, \text{ em que: } m_i = m + t_i \begin{cases} i=1, 2, \dots, I \\ j=1, 2, \dots, J \end{cases}$$

Considerando-se, para obtenção dos componentes de variância do modelo aleatório, que os efeitos aleatórios do modelo sejam não correlacionados, apresentam-se a seguir diferentes possibilidades para a análise de variância:

a) Análise de variância (ANOVA) em ensaio casualizado em blocos com t_i : aleatório e b_j : aleatório:

m : média geral (fixa), tal que $E(m) = m$ e $E(m^2) = m^2$;

t_i : efeito aleatório do i -ésimo tratamento, em que $t_i \sim NID(0, \sigma_t^2)$, ou seja, normal e independentemente distribuído, com média zero e variância σ_t^2 , logo $E(t_i) = 0$ e $E(t_i^2) = \sigma_t^2$;

b_j : efeito aleatório do j -ésimo bloco, em que $b_j \sim NID(0, \sigma_b^2)$, ou seja, normal e independentemente distribuído, com média zero e variância σ_b^2 , logo $E(b_j) = 0$ e $E(b_j^2) = \sigma_b^2$;

e_{ij} : é o erro experimental, considerado aleatório, normal e independentemente distribuído, isto é, $e_{ij} \sim NID(0, \sigma^2)$, logo $E(e_{ij}) = 0$ e $E(e_{ij}^2) = \sigma^2$.

Tabela 4.2.2. ANOVA casualizado em blocos (t_i : aleatório e b_j : aleatório).

F. V.	G.L.	S.Q.	Q.M.	E(Q.M.)	F
Trat.	$I - 1$	$\frac{1}{J} \sum_{i=1}^I T_i^2 - C$	$\frac{SQ_{\text{TRAT}}}{I - 1}$	$\sigma^2 + J\sigma_t^2$	$\frac{QM_{\text{TRAT}}}{QM_{\text{RES}}}$
Blocos	$J - 1$	$\frac{1}{I} \sum_{j=1}^J B_j^2 - C$	$\frac{SQ_{\text{BLOCOS}}}{J - 1}$	$\sigma^2 + I\sigma_b^2$	$\frac{QM_{\text{BLOCOS}}}{QM_{\text{RES}}}$
Resíduo	$(I - 1)(J - 1)$	Diferença	$\frac{SQ_{\text{RES}}}{(I - 1)(J - 1)}$	σ^2	
Total	$IJ - 1$	$\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C$			

b) ANOVA para o ensaio casualizado em blocos com t_i : fixo e b_j : fixo:

t_i : efeito fixo do i -ésimo tratamento com $E(t_i) = t_i$, $E(t_i^2) = t_i^2$ e $\sum_{i=1}^I t_i = 0$;

b_j : efeito fixo do j -ésimo bloco com $E(b_j) = b_j$, $E(b_j^2) = b_j^2$ e $\sum_{j=1}^J b_j = 0$;

e_{ij} : é o erro experimental, aleatório, normal e independentemente distribuído, ou seja, $e_{ij} \sim NID(0, \sigma^2)$, logo $E(e_{ij}) = 0$, $E(e_{ij}^2) = \sigma^2$ e $Cov(e_{ij}, e_{ij'}) = 0$ para $(i \neq i' \text{ e } j \neq j')$.

Tabela 4.2.3. ANOVA casualizado em blocos (t_i : fixo e b_j : fixo).

F. V.	G.L.	S.Q.	Q.M.	E(Q.M.)	F
Trat.	$I - 1$	$\frac{1}{J} \sum_{i=1}^I T_i^2 - C$	$\frac{SQ_{TRAT}}{I-1}$	$\sigma^2 + J\phi_t$	$\frac{QM_{TRAT}}{QM_{RES}}$
Blocos	$J - 1$	$\frac{1}{I} \sum_{j=1}^J B_j^2 - C$	$\frac{SQ_{BLOCOS}}{J-1}$	$\sigma^2 + I\phi_b$	$\frac{QM_{BLOCOS}}{QM_{RES}}$
Resíduo	$(I - 1)(J - 1)$	Diferença	$\frac{SQ_{RES}}{(I-1)(J-1)}$	σ^2	
Total	$IJ - 1$	$\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C$			

$$\phi_t = \frac{\sum_{i=1}^I t_i^2}{I-1} \quad \phi_b = \frac{\sum_{j=1}^J b_j^2}{J-1}$$

c) ANOVA para o experimento casualizado em blocos com t_i : fixo e b_j : aleatório:

t_i : efeito fixo do i -ésimo tratamento com $E(t_i) = t_i$, $E(t_i^2) = t_i^2$ e $\sum_{i=1}^I t_i = 0$;

b_j : efeito aleatório do j -ésimo bloco [$b_j \sim NID(0, \sigma_b^2)$], normal e independentemente distribuído, com média zero e variância σ_b^2 , logo $E(b_j) = 0$ e $E(b_j^2) = \sigma_b^2$.

Tabela 4.2.4. ANOVA casualizado em blocos (t_i : fixo e b_j : aleatório).

F. V.	G.L.	S.Q.	Q.M.	E(Q.M.)	F
Trat.	$I - 1$	$\frac{1}{J} \sum_{i=1}^I T_i^2 - C$	$\frac{SQ_{TRAT}}{I-1}$	$\sigma^2 + J\phi_t$	$\frac{QM_{TRAT}}{QM_{RES}}$
Blocos	$J - 1$	$\frac{1}{I} \sum_{j=1}^J B_j^2 - C$	$\frac{SQ_{BLOCOS}}{J-1}$	$\sigma^2 + I\sigma_b^2$	$\frac{QM_{BLOCOS}}{QM_{RES}}$
Resíduo	$(I - 1)(J - 1)$	Diferença	$\frac{SQ_{RES}}{(I-1)(J-1)}$	σ^2	
Total	$IJ - 1$	$\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C$			

$$\phi_t = \frac{\sum_{i=1}^I t_i^2}{I-1}$$

d) ANOVA para o delineamento casualizado em blocos com t_i aleatório e b_j fixo:

t_i : efeito aleatório do i -ésimo tratamento, em que $t_i \sim NID(0, \sigma_t^2)$, ou seja, normal e independentemente distribuído, com média zero e variância σ_t^2 , logo $E(t_i) = 0$ e $E(t_i^2) = \sigma_t^2$;

b_j : efeito fixo do j -ésimo bloco com $E(b_j) = b_j$, $E(b_j^2) = b_j^2$ e $\sum_{j=1}^J b_j = 0$.

Tabela 4.2.5. ANOVA casualizado em blocos (t_i : aleatório e b_j : fixo).

F. V.	G.L.	S.Q.	Q.M.	E(Q.M.)	F
Trat.	$I - 1$	$\frac{1}{J} \sum_{i=1}^I T_i^2 - C$	$\frac{SQ_{TRAT}}{I - 1}$	$\sigma^2 + J\sigma_t^2$	$\frac{QM_{TRAT}}{QM_{RES}}$
Blocos	$J - 1$	$\frac{1}{I} \sum_{j=1}^J B_j^2 - C$	$\frac{SQ_{BLOCOS}}{J - 1}$	$\sigma^2 + I\phi_b$	$\frac{QM_{BLOCOS}}{QM_{RES}}$
Resíduo	$(I - 1)(J - 1)$	Diferença	$\frac{SQ_{RES}}{(I - 1)(J - 1)}$	σ^2	
Total	$IJ - 1$	$\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C$			

$$\phi_b = \frac{\sum_{j=1}^J b_j^2}{J - 1}$$

Exercício 4.2.1. Considerar o experimento, casualizado em blocos, concernente ao estudo de competição de quatro espécies florestais (E1, E2, E3 e E4), sabendo-se que os dados correspondem ao volume de madeira por parcela na data correspondente à rotação economicamente aconselhável. Nesse contexto, efetuar a análise de variância, verificando a normalidade dos resíduos e aditividade dos efeitos pelo teste de Tukey.

Tabela 4.2.6. Volume de madeira em m³ por parcela (Dados simulados).

Blocos	E1	E2	E3	E4	Total
1	43,3	59,7	56,7	49,0	208,7
2	40,4	52,8	56,7	46,7	196,6
3	42,5	63,0	60,5	54,0	220,0
4	38,5	51,3	70,0	46,4	206,2
5	44,3	56,3	66,2	51,8	218,6
Total	209,0	283,1	310,1	247,9	1.050,1

a) Análise de variância

$$SQ_{TOTAL} = 43,3^2 + 40,4^2 + \dots + 46,4^2 + 51,8^2 - C$$

$$C = \frac{1.050,1^2}{4 \times 5} = 55.135,5005$$

$$SQ_{TOTAL} = 56.586,91 - C = 1.451,4095$$

$$SQ_{BLOCOS} = \frac{208,7^2 + 196,6^2 + 220,0^2 + 206,2^2 + 218,6^2}{4} - C = 92,4120$$

$$SQ_{TRAT} = \frac{209,0^2 + 283,1^2 + 310,1^2 + 247,9^2}{5} - C = 1.153,1055$$

$$SQ_{RES} = SQ_{TOTAL} - SQ_{BLOCOS} - SQ_{TRAT} = 205,8920$$

Tabela 4.2.7. ANOVA do ensaio casualizado em blocos.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	4	92,4120	23,1030	1,35
Espécies	3	1.153,1055	384,3685	22,40 **
Resíduo	12	205,8920	17,1577	
Total	19	1.451,4095		

Dado que $F_{3; 12; 0,01} = 5,95$, considerando o nível de significância de $\alpha = 0,01$, conclui-se pela rejeição de $H_0 : m_1 = m_2 = m_3 = m_4$.

b) Testes de normalidade

Os desvios (e_{ij}) dos dados do exercício 4.2.1 — isto é, $e_{ij} = x_{ij} - m - t_i - b_j$ (tabela 4.2.8) —, são importantes para verificar se a distribuição dos resíduos do modelo linear segue a distribuição normal, em que $\hat{m} = 52,505$.

Tabela 4.2.8. Desvios ($e_{ij} = x_{ij} - m - t_i - b_j$).

Trat(i)	Blocos (j)	x_{ij}	$\hat{m}_i$	$\hat{t}_i = \hat{m}_i - \hat{m}$	$\hat{m}_j$	$\hat{b}_j = \hat{m}_j - \hat{m}$	e_{ij}
E1	1	43,3	41,80	-10,705	52,175	-0,330	1,830
E1	2	40,4	41,80	-10,705	49,150	-3,355	1,955
E1	3	42,5	41,80	-10,705	55,000	2,495	-1,795
E1	4	38,5	41,80	-10,705	51,550	-0,955	-2,345
E1	5	44,3	41,80	-10,705	54,650	2,145	0,355
E2	1	59,7	56,62	4,115	52,175	-0,330	3,410
E2	2	52,8	56,62	4,115	49,150	-3,355	-0,465
E2	3	63,0	56,62	4,115	55,000	2,495	3,885
E2	4	51,3	56,62	4,115	51,550	-0,955	-4,365
E2	5	56,3	56,62	4,115	54,650	2,145	-2,465
E3	1	56,7	62,02	9,515	52,175	-0,330	-4,990
E3	2	56,7	62,02	9,515	49,150	-3,355	-1,965
E3	3	60,5	62,02	9,515	55,000	2,495	-4,015
E3	4	70,0	62,02	9,515	51,550	-0,955	8,935
E3	5	66,2	62,02	9,515	54,650	2,145	2,035
E4	1	49,0	49,58	-2,925	52,175	-0,330	-0,250
E4	2	46,7	49,58	-2,925	49,150	-3,355	0,475
E4	3	54,0	49,58	-2,925	55,000	2,495	1,925
E4	4	46,4	49,58	-2,925	51,550	-0,955	-2,225
E4	5	51,8	49,58	-2,925	54,650	2,145	0,075

Os dados das observações do exercício 4.2.1 foram analisados pelo programa SAS, **Prog5a**, em anexo no apêndice B, para aplicar os testes de Shapiro-Wilk, Kolmogorov, Cramér-Von Mises e Anderson-Darling, no intuito de verificar se a distribuição dos resíduos, relativa ao modelo linear utilizado, segue a distribuição normal. Os resultados, apresentados na tabela 4.2.9, mostram que os resíduos podem ser considerados normalmente distribuídos.

O programa SAS, **Prog5b**, em anexo no apêndice B, foi usado para realizar os testes de Kolmogorov-Smirnov, Cramér-Von Mises, Anderson-Darling e Shapiro-Wilk para os valores dos desvios $e_{ij} = x_{ij} - m - t_i - b_j$, (tabela 4.2.8), ao invés das observações, e assim verificar se a distribuição dos resíduos do modelo linear utilizado segue a distribuição normal. Os resultados foram os mesmos que os apresentados na tabela 4.2.9, isto é, quando foi usado o programa **Prog5a**.

Tabela 4.2.9. Resultados dos testes de normalidade no SAS.

Testes de Normalidade				
Testes	Estatísticas		p-valor	
Shapiro-Wilk	W	0,944816	Pr < W	0,2952
Kolmogorov-Smirnov	D	0,118225	Pr > D	> 0,1500
Cramér-Von Mises	W-Sq	0,040894	Pr > W-Sq	0,2500
Anderson-Darling	A-Sq	0,325163	Pr > A-Sq	0,2500

c) Teste de não aditividade de Tukey

De acordo com Steel e Torrie (1960), os seguintes passos são necessários para o teste da não aditividade dos efeitos:

c1) Obter os desvios das médias dos tratamentos $(\bar{x}_i - \bar{x}_{..})$ e dos blocos $(\bar{x}_{.j} - \bar{x}_{..})$, em a relação à média geral $(\bar{x}_{..} = 52,505)$;

c2) Calcular para cada bloco $Q_j = \sum_{i=1}^I (\bar{x}_i - \bar{x}_{..}) X_{ij}$, tal que X_{ij} são os valores da variável resposta ($i = 1, \dots, I$ e $j = 1, \dots, J$), em que $Q_j = \sum_{i=1}^I (\bar{x}_i - \bar{x}_{..}) X_{ij}$ é equivalente ao cálculo do numerador do coeficiente de uma regressão linear simples;

c3) Obter $Q = \sum_{j=1}^J (\bar{x}_{.j} - \bar{x}_{..}) Q_j$, ou seja, $Q = \sum_{i,j} (\bar{x}_i - \bar{x}_{..})(\bar{x}_{.j} - \bar{x}_{..}) X_{ij}$;

c4) E finalmente obter:

$$VP = \frac{Q^2}{\sum_{i=1}^I (\bar{x}_i - \bar{x}_{..})^2 \times \sum_{j=1}^J (\bar{x}_{.j} - \bar{x}_{..})^2}$$

Tabela 4.2.10. Cálculos para o teste de não aditividade de Tukey.

Blocos	E1	E2	E3	E4	$\bar{x}_j$	$(\bar{x}_j - \bar{x}_{..})$	$Q_j = \sum_{i=1}^I (\bar{x}_i - \bar{x}_{..}) X_{ij}$
1	43,3	59,7	56,7	49,0	52,175	-0,330	178,3145
2	40,4	52,8	56,7	46,7	49,15	-3,355	187,6930
3	42,5	63,0	60,5	54,0	55,00	2,495	221,9900
4	38,5	51,3	70,0	46,4	51,55	-0,955	329,2870
5	44,3	56,3	66,2	51,8	54,65	2,145	235,8210
$\bar{x}_i$	41,8	56,62	62,02	49,58	52,505		
$(\bar{x}_i - \bar{x}_{..})$	-10,705	4,115	9,515	-2,925		0	

Então:

$$\sum_{i=1}^I (\bar{x}_i - \bar{x}_{..})^2 = (-10,705)^2 + 4,115^2 + 9,515^2 + 9,515^2 + (-2,925)^2 = 230,6211$$

$$\sum_{j=1}^J (\bar{x}_j - \bar{x}_{..})^2 = (-0,33)^2 + (-3,355)^2 + 2,495^2 + (-0,955)^2 + 2,145^2 = 23,103$$

$$Q = \sum_{j=1}^J (\bar{x}_j - \bar{x}_{..}) Q_j = -0,33 \times 178,3145 - 3,355 \times 187,693 + 2,495 \times 221,99 - 0,955 \times 329,287 + 2,145 \times 235,821$$

$$Q = \sum_{j=1}^J (\bar{x}_j - \bar{x}_{..}) Q_j = 56,67821$$

$$VP = \frac{Q^2}{\sum_{i=1}^I (\bar{x}_i - \bar{x}_{..})^2 \times \sum_{j=1}^J (\bar{x}_j - \bar{x}_{..})^2} = \frac{56,67821^2}{230,6211 \times 23,103} = 0,602927$$

Tabela 4.2.11. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	4	92,412000	23,103000	1,24
Espécies	3	1.153,105500	384,368500	20,60 **
VP	1	0,602927	0,602927	0,03
Resíduo	11	205,289073	18,662600	
Total	19	1.451,409500		

Verifica-se, na tabela 4.2.11, que o valor da estatística do teste é igual a $F_c = 0,03$, apresentando $v_1 = 1$ e $v_2 = 11$ graus de liberdade. Sendo o p - valor correspondendo a 0,8606 (tabela 4.2.12), conclui-se que a hipótese da nulidade não deve ser rejeitada, havendo evidência para afirmar que os efeitos atribuídos ao modelo linear utilizado são aditivos.

Na tabela 4.2.12, está a análise de não aditividade de Tukey (1949a) realizada pelo SAS, **Prog7**, em anexo no apêndice B:

Tabela 4.2.12. Análise de não aditividade de Tukey pelo SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	4	92,412000	23,103000	1,24	0,3506
Espécies	3	1153,105500	384,368500	20,60	< 0,0001
VP	1	0,602927	0,602927	0,03	0,8606
Resíduo	11	205,289073	18,662643		
Total	19	1451,409500			

Por outro lado, no caso de rejeição da hipótese nula de não aditividade, alguma forma de transformação dos dados deve ser testada para corrigir o problema, o que não é o caso do exemplo apresentado.

Exercício 4.2.2. Considerar um ensaio casualizado em blocos, apresentado em Pimentel Gomes e Garcia (2002), sobre o comportamento de sete progênies de *Eucalyptus grandis*, conforme tabela 4.2.13, em que os valores da produção de madeira em m³ ha⁻¹ para as sete progênies são apresentados.

Tabela 4.2.13. Produção de madeira (m³ ha⁻¹) para as sete progênies.

Progênies	P1	P2	P3	P4	P5	P6	P7	Totais
Bloco 1	358	284	273	284	258	249	318	2.024,0
Bloco 2	380	249	222	242	263	217	312	1.885,0
Bloco 3	353	259	236	266	242	267	327	1.950,0
Bloco 4	360	242	226	252	231	220	319	1.850,0
Total	1.451,0	1.034,0	957,0	1.044,0	994,0	953,0	1.276,0	7.709,0

$$C = \frac{7.709,0^2}{7 \times 4} = 2.122.452,8929$$

$$SQ_{TOTAL} = 358^2 + 380^2 + \dots + 327^2 + 319^2 - C = 60.282,1071$$

$$SQ_{BLOCOS} = \frac{2.024,0^2 + 1.885,0^2 + 1.950,0^2 + 1.850,0^2}{7} - C$$

$$SQ_{BLOCOS} = \frac{14.874.801,0}{7} - C = 2.518,6786$$

$$SQ_{TRAT} = \frac{1.451,0^2 + 1.034,0^2 + \dots + 1.276,0^2}{4} - C = 53.737,8571$$

$$SQ_{TRAT} = \frac{8.704.763,0}{4} - C = 53.737,8571$$

$$SQ_{RES} = SQ_{TOTAL} - SQ_{BLOCOS} - SQ_{TRAT}$$

$$SQ_{RES} = 60.282,1071 - 2.518,6786 - 53.737,8571 = 4.025,5714$$

Tabela 4.2.14. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	3	2.518,6786	839,5595	3,75
Progenies	6	53.737,8571	8.956,3095	40,05 **
Resíduo	18	4.025,5714	223,6429	
Total	27	60.282,1071		

Dado que $F_{6;18;0,01} = 4,01$, conclui-se, ao nível de significância $\alpha = 0,01$, pela rejeição da hipótese de nulidade: $H_0 : m_1 = m_2 = \dots = m_7$.

4.3. EFETIVIDADE DO DELINEAMENTO CASUALIZADO EM BLOCOS

Quando utilizado o delineamento casualizado em blocos, é importante determinar a sua efetividade em relação ao delineamento inteiramente ao acaso. O objetivo do teste da efetividade dos blocos é avaliar o uso de blocos no aumento ou não da precisão do experimento. De acordo com Neter *et al.* (1990), a efetividade pode ser obtida pela expressão:

$$EB = \frac{(J-1)QM_{BLOCOS} + J(I-1)QM_{RES}}{(IJ-1)QM_{RES}}$$

A estatística EB indica quantas repetições seriam necessárias no delineamento inteiramente ao acaso quando comparado com delineamento casualizado em blocos.

Quando o valor de EB for menor que 1 (um) indica perda de precisão pelo uso de blocos.

Exercício 4.3.1. O experimento casualizado em blocos do exercício 4.2.1 apresenta o estudo de competição de quatro espécies florestais (E1, E2, E3 e E4). Na tabela 4.3.1, está o resultado da análise de variância. Nesse contexto, estimar a efetividade em relação ao delineamento inteiramente ao acaso.

Tabela 4.3.1. ANOVA do ensaio casualizado em blocos.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	4	92,4120	23,1030	1,35
Espécies	3	1.153,1055	384,3685	22,40 **
Resíduo	12	205,8920	17,1577	
Total	19	1.451,4095		

$$EB = \frac{(J-1)QM_{BLOCOS} + J(I-1)QM_{RES}}{(IJ-1)QM_{RES}} \quad EB = \frac{(5-1) \times 23,1030 + 5(4-1)17,1577}{(4 \times 5 - 1) \times 17,1577} = 1,07$$

A estatística EB indica uma relação inerente às repetições que seriam necessárias no delineamento inteiramente ao acaso quando comparado com delineamento casualizado em blocos.

O valor $EB = 1,07$ indica que, considerando cada bloco seriam necessárias $1,07 \times J$ repetições no delineamento inteiramente ao acaso. Assim sendo, o uso de blocos foi efetivo. Por outro lado, considerando que $1,07$ está muito próximo de 1 (um), pode-se considerar que a efetividade no uso de blocos apresenta a mesma eficácia caso fosse aplicado o delineamento inteiramente ao acaso.

Silva e Silva (2002) comentam que no caso em que EB for menor do que 1 (um), existe a indicação de que há perda de precisão pelo uso do delineamento casualizado em blocos. Por outro lado, mencionam que, apesar de violar o princípio da casualização, o que se poderia fazer, quando EB for menor do que 1 (um), é efetuar a análise de variância como se fosse o delineamento inteiramente ao acaso.

4.4. EXPERIMENTOS CASUALIZADOS EM BLOCOS COM TRATAMENTOS COMUNS

O planejamento de ensaios com o número elevado de tratamentos é comum na experimentação florestal, o que dificulta a utilização do delineamento casualizado em blocos, haja vista a necessidade de estabelecer blocos com tamanhos excessivos, o que pode comprometer a sua homogeneidade.

Em contrapartida, em um experimento com grande número de procedências de uma espécie florestal, é possível agrupá-las, tal que, em cada grupo, sejam adicionadas procedências consideradas importantes e já estudadas. Completando os grupos, pode ser utilizado o delineamento em blocos ao acaso. De acordo com Pimentel Gomes (2000), a análise conjunta é relativamente fácil e corresponde ao uso de um delineamento em blocos incompletos, de grande flexibilidade e eficiência.

Campos e Piedade (1992) sugerem os seguintes passos para realizar o planejamento experimental:

- a) Reunir os tratamentos, preferencialmente, em grupos de mesmos tamanhos;
- b) Constituir cada grupo em um experimento casualizado em blocos;
- c) Tomar alguns tratamentos “adicionais” (exaustivamente testados ou com comportamentos conhecidos, que integrarão todos os grupos — denominados “Tratamentos comuns”). Ressalte-se que os originais, específicos de cada grupo, são denominados “Tratamentos regulares”;
- d) Realizar, uma vez coletados os dados de cada experimento, separadamente, as análises estatísticas de acordo com o delineamento casualizado em blocos;
- e) Realizar a análise conjunta de todos os tratamentos, em que os tratamentos comuns sejam considerados o “elo de ligação”;
- f) Possibilitar, por meio de análises, quatro tipos de comparações entre dois tratamentos:
 - f1) Entre dois tratamentos regulares de um mesmo grupo;
 - f2) Entre dois tratamentos regulares de grupos distintos;

f3) Entre um tratamento regular e um tratamento comum;

f4) Entre dois tratamentos comuns.

4.4.1. Análise de variância de ensaios casualizados em blocos com tratamentos comuns

Considerar o exemplo de um experimento com 18 procedências de uma determinada espécie florestal, em que se pretende compará-las com duas procedências, já experimentadas e consagradas como de excelente potencial, denominadas de tratamentos adicionais PC1 e PC2 (Tratamentos comuns). As 18 procedências (Tratamentos regulares) são reunidas em três grupos, tal que em cada agrupamento foram adicionadas as procedências PC1 e PC2, como segue:

Grupo1	P1	P2	P3	P4	P5	P6	PC1	PC2
Grupo2	P7	P8	P9	P10	P11	P12	PC1	PC2
Grupo3	P13	P14	P15	P16	P17	P18	PC1	PC2

Considerando o uso de quatro blocos, com cada grupo com oito tratamentos, será formado um ensaio casualizado em blocos, o qual apresenta o seguinte esquema de análise de variância:

F.V.	G.L.
Blocos	3
Tratamentos	7
Resíduo	21
Total	31

Logo, a estrutura da análise de variância conjunta será:

F.V.	G.L.
Experimentos (grupos)	$n_g - 1 = 3 - 1 = 2$
Blocos dentro de experimento	$n_g (J - 1) = 3(4 - 1) = 9$
Tratamentos ajustados	$n_g n_r + n_c - 1 = 3 \times 6 + 2 - 1 = 19$
Interação Trat. comuns $\times$ Experimentos	$(n_c - 1)(n_g - 1) = (2 - 1)(3 - 1) = 2$
Resíduo	$n_g (J - 1)(n_r + n_c - 1) = 3(4 - 1)(6 + 2 - 1) = 63$
Total	$(n_r + n_c) J n_g - 1 = (6 + 2) \times 4 \times 3 - 1 = 95$

n_g : número de grupos;

n_c : número de tratamentos comuns;

n_r : número de tratamentos regulares por grupo;

J : número de blocos

Nessa análise há quatro alternativas de comparações entre tratamentos:

a) Comparação entre dois tratamentos comuns:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2}{n_g J} \hat{\sigma}^2 \quad \hat{\sigma}^2: \text{quadrado médio da análise conjunta}$$

b) Comparação entre dois tratamentos regulares de um mesmo experimento:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2}{J} \hat{\sigma}^2$$

c) Comparação entre dois tratamentos regulares de diferentes experimentos:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2}{J} \left(1 + \frac{1}{n_c}\right) \hat{\sigma}^2$$

d) Comparação entre um tratamento comum e um regular:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{1}{J} \left(1 + \frac{1}{n_c} + \frac{1}{n_g} - \frac{1}{n_c n_g}\right) \hat{\sigma}^2$$

Para a realização dos testes de comparações múltiplas, as médias dos tratamentos regulares devem ser ajustadas em relação aos tratamentos comuns. O ajuste é feito da seguinte forma:

$$\hat{m}_{ik}(Aj.) = \hat{m}_{ik} - \hat{t}_k \quad (k = 1, \dots, n_g)$$

Em que $\hat{t}_k$ é a diferença estimada entre a média dos tratamentos comuns no respectivo experimento e a média geral dos tratamentos comuns.

No caso dos tratamentos comuns usa-se, sem nenhum ajuste, a média aritmética.

Exercício 4.4.1.1. Considerar os dados simulados de produção madeireira em m³ por parcela (tabelas 4.4.1.1, 4.4.1.2 e 4.4.1.3), referentes à produção de 20 procedências de uma espécie florestal, tal que as procedências PC1 e PC2 são os tratamentos comuns, enquanto as procedências P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P14, P15, P16, P17 e P18 são as denominadas procedências regulares. Nesse contexto, efetuar a análise estatística.

Tabela 4.4.1.1. Produção de madeira (m³ por parcela) das procedências do grupo 1.

Blocos	G1: Procedências								Total
	PC1	PC2	P1	P2	P3	P4	P5	P6	
1	85,5	84,6	86,3	83,8	87,9	84,6	82,2	89,5	684,4
2	83,0	83,8	87,1	83,0	88,7	86,3	83,8	91,2	686,9
3	84,6	85,5	85,5	84,6	91,2	85,5	83,0	88,7	688,6
4	86,3	83,5	86,3	82,2	89,50	87,1	82,2	90,3	687,4
Total	339,4	337,4	345,2	333,6	357,3	343,5	331,2	359,7	2.747,3

Tabela 4.4.1.2. Produção de madeira (m³ por parcela) das procedências do grupo 2.

Blocos	G2: Procedências								Total
	PC1	PC2	P7	P8	P9	P10	P11	P12	
1	86,3	83,0	83,8	82,2	84,6	90,3	92,0	87,1	689,3
2	83,0	84,6	86,3	80,6	85,5	92,0	90,3	86,3	688,6
3	83,8	86,3	84,6	81,4	86,3	89,5	91,2	87,9	691,0
4	84,6	83,8	82,2	83,8	83,0	92,8	93,6	85,5	689,3
Total	337,7	337,7	336,9	328,0	339,4	364,6	367,1	346,8	2.758,2

Tabela 4.4.1.3. Produção de madeira (m³ por parcela) das procedências do grupo 3.

Blocos	G3: Procedências								Total
	PC1	PC2	P13	P14	P15	P16	P17	P18	
1	88,7	85,5	91,2	86,3	83,8	89,5	91,2	92,8	709,0
2	87,1	87,9	91,2	85,5	86,3	90,3	93,6	91,2	713,1
3	85,5	84,6	92,0	87,1	87,9	91,2	92,8	92,0	713,1
4	86,3	87,1	89,5	87,9	83,0	87,9	94,4	91,2	707,3
Total	347,6	345,1	363,9	346,8	341,0	358,9	372,0	367,2	2.842,5

O contexto da estrutura experimental mostra um conjunto de três ensaios, correspondente a um experimento em blocos incompletos com vinte tratamentos, ou seja, 18 novas procedências (tratamentos regulares) e as duas procedências já conhecidas, consideradas padrões (tratamentos comuns). Nota-se que, considerando dois tratamentos regulares (i, j), o número de vezes que eles aparecem quando forem do mesmo grupo, no mesmo bloco, será igual a quatro. Os tratamentos regulares P13 e P14 aparecem quatro vezes no grupo ou experimento 3. Contudo, quando comparados tratamentos regulares de grupos diferentes, há o resultado $\lambda(i, j) = 0$, por exemplo: P8 e P13. Quando considerados os tratamentos comuns PC1 e PC2, ocorre $\lambda(i, j) = 12$. Quando considerado um tratamento regular (i) qualquer e um tratamento comum (j), por exemplo P13 e PC1, então $\lambda(i, j) = 4$. Sendo as análises de variâncias individuais:

Análise de variância do grupo 1				
F.V.	G.L.	S.Q.	Q.M.	F
Blocos	3	1,1709	0,3903	
Procedências	7	189,7572	27,1082	21,60 **
Resíduo	21	26,3515	1,2548	
Total	31	217,2797		

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	1,1709375	0,3903125	0,31	0,8172
Procedências	7	189,7571875	27,1081696	21,60	< 0,0001
Resíduo	21	26,3515625	1,2548363		
Total	31	217,2796875			

Análise de variância do grupo2				
F.V.	G.L.	S.Q.	Q.M.	F
Blocos	3	0,3913	0,1304	
Procedências	7	341,9888	48,8555	21,44
Resíduo	21	47,8588	2,2790	
Total	31	390,2388		

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	0,3912500	0,1304167	0,06	0,9815
Procedências	7	341,9887500	48,8555357	21,44	< 0,0001
Resíduo	21	47,8587500	2,2789881		
Total	31	390,2387500			

Análise de variância do grupo3				
F.V.	G.L.	S.Q.	Q.M.	F
Blocos	3	3,2434	1,0811	
Procedências	7	236,8722	33,8389	16,10 **
Resíduo	21	44,1491	2,1023	
Total	31	284,2647		

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	3,2434375	1,0811458	0,51	0,6769
Procedências	7	236,8721875	33,8388839	16,10	< 0,0001
Resíduo	21	44,1490625	2,1023363		
Total	31	284,2646875			

Análise de variância conjunta:

$$C = \frac{(85,5 + \dots + 90,3 + 86,3 + \dots + 85,5 + 88,7 + \dots + 91,2)^2}{96} = \frac{8.3480^2}{96} = 725.928,1667$$

$$SQ_{TOTAL} = 85,5^2 + \dots + 90,3^2 + 86,3^2 + \dots + 85,5^2 + 88,7^2 + \dots + 91,2^2 - C$$

$$SQ_{TOTAL} = 726.989,62 - C = 1.061,4533$$

$$SQ_{GRUPOS} = \frac{2.747,3^2 + 2.758,2^2 + 2.842,5^2}{32} - C = 169,6702$$

$$SQ_{BLOCOS \text{ D. GRUPO}} = SQ_{BLOCOS (G1)} + SQ_{BLOCOS (G2)} + SQ_{BLOCO (G3)}$$

$$SQ_{BLOCOS \text{ D. GRUPO}} = 1,1709 + 0,3913 + 3,2434 = 4,8056$$

$$SQ_{RESÍDUO} = SQ_{RESÍDUO(G1)} + SQ_{RESÍDUO(G2)} + SQ_{RESÍDUO(G3)}$$

$$SQ_{RESÍDUO} = 26,3515 + 47,8588 + 44,1491 = 118,3594$$

Quadro auxiliar, considerando os tratamentos comuns PC1 e PC2:

Trat. (PC)	G1	G2	G3	Totais
PC1	339,4	337,7	347,6	1.024,7
PC2	337,4	337,7	345,1	1.020,2
Totais	676,8	675,4	692,7	2.044,9

Soma de quadrados entre as procedências comuns (SQ_{PC}):

$$SQ_{PC} = \frac{1.024,7^2 + 1.020,2^2}{12} - \frac{2.044,9^2}{24} = 0,8438$$

Soma de quadrados das procedências comuns dentro dos grupos (SQ_{GPC}):

$$SQ_{GPC} = \frac{676,8^2 + 675,4^2 + 692,7^2}{8} - \frac{2.044,9^2}{24} = 23,0858$$

Soma de quadrados, considerando as combinações das procedências comuns com os grupos ($SQ_{PC,GRUPO}$):

$$SQ_{PC,GRUPO} = \frac{339,4^2 + 337,7^2 + 347,6^2 + 337,4^2 + 337,7^2 + 345,1^2}{4} - \frac{2.044,9^2}{24} = 24,3671$$

Soma de quadrados da interação entre as procedências comuns e grupos ($SQ_{PC \times GRUPO}$):

$$SQ_{PC \times GRUPO} = SQ_{PC,GRUPO} - SQ_{PC} - SQ_{GPC}$$

$$SQ_{PC \times GRUPO} = 24,3671 - 0,8438 - 23,0858 = 0,4375$$

Soma de quadrados ajustada entre as procedências ($SQ_{PROC(Aj.)}$):

$$SQ_{PROC(Aj.)} = SQ_{TOTAL} - (SQ_{GRUPOS} + SQ_{BLOCOS \text{ D. GRUPO}} + SQ_{PC \times GRUPO} + SQ_{RESÍDUO})$$

$$SQ_{PROC(Aj.)} = 1.061,4533 - (169,6702 + 4,8056 + 0,4375 + 118,3594) = 768,1806$$

É importante observar, na análise de variância conjunta (tabela 4.4.1.4) que a interação entre os tratamentos comuns e os grupos (experimentos), Grupos $\times$ Proc(C), é não significativa, o que indica a existência de uma identidade de comportamento dos tratamentos comuns (PC1 e PC2) com todos os experimentos, confirmando a sua função de “elo de ligação”.

Tabela 4.4.1.4. ANOVA conjunta.

F.V.	G.L.	S.Q.	Q.M.	F
Grupos	2	169,6702	84,8351	
Blocos d. grupos	9	4,8056	0,5340	
Procedências (Aj.)	19	768,1806	40,4306	21,52 **
Proc(C)×Grupo	2	0,4375	0,2188	0,12
Resíduo	63	118,3594	1,8787	
Total	95	1.061,4533		

Seja a análise realizada pelo programa SAS, **Prog16**, em anexo no apêndice B:

Tabela 4.4.1.5. ANOVA conjunta no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Grupos	2	169,6702083	84,8351042	45,16	< 0,0001
Blocos d. (Grupos)	9	4,8056250	0,5339583	0,28	0,9767
Procedências (Aj.)	19	768,1806250	40,4305592	21,52	< 0,0001
Grupos×Proc (C)	2	0,4375000	0,2187500	0,12	0,8903
Resíduo	63	118,359375	1,878720		
Total	95	1061,453333			

Para a realização dos testes de comparações múltiplas, as médias dos tratamentos regulares devem ser ajustadas em relação aos tratamentos comuns. O ajuste é feito da seguinte forma:

$$\hat{m}_{ik}(Aj.) = \hat{m}_{ik} - \hat{t}_k$$

Em que $\hat{t}_k$ é a diferença entre a média dos tratamentos comuns do j -ésimo experimento versus a média geral, considerando todos os tratamentos comuns, dado que:

Trat. (PC)	G1	G2	G3	Totais
PC1	339,4	337,7	347,6	1.024,7
PC2	337,4	337,7	345,1	1.020,2
Médias	84,6000	84,4250	86,5875	85,2042

Logo, considerando os três experimentos, tem-se:

$$\hat{t}_1 = 84,60 - 85,2042 = -0,6042$$

$$\hat{t}_2 = 84,425 - 85,2042 = -0,7792$$

$$\hat{t}_3 = 86,5875 - 85,2042 = 1,3833$$

a) Tratamentos comuns:

Procedências	Média não ajustada	Média ajustada
PC1	85,3917	85,3917
PC2	85,0167	85,0167

b) Médias ajustadas dos tratamentos regulares:

Experimento ou grupo1		
Procedências	Média não ajustada	Média ajustada [$\hat{m}_{i1} - (-0,6042)$]
P1	86,300	86,9042
P2	83,400	84,0042
P3	89,325	89,9292
P4	85,875	86,4792
P5	82,800	83,4042
P6	89,925	90,5292

Experimento ou grupo2		
Procedências	Média não ajustada	Média ajustada [$\hat{m}_{i2} - (-0,7792)$]
P7	84,225	85,0042
P8	82,000	82,7792
P9	84,850	85,6292
P10	91,150	91,9292
P11	91,775	92,5542
P12	86,700	87,4792

Experimento ou grupo3		
Procedências	Média não ajustada	Média ajustada [$\hat{m}_{i3} - (1,3833)$]
P13	90,975	89,5917
P14	86,700	85,3167
P15	85,250	83,8667
P16	89,725	88,3417
P17	93,000	91,6167
P18	91,800	90,4167

J : número de blocos por experimento;

n_g : número de grupos ou de experimentos;

$\hat{\sigma}^2$: quadrado médio da análise conjunta.

Existem as seguintes quatro alternativas de comparações entre duas médias:

a) Na comparação entre dois tratamentos comuns

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2}{J \times n_g} \hat{\sigma}^2 = \frac{2}{4 \times 3} \times 1,8787 = 0,3131$$

O teste de Tukey para comparar as duas procedências comuns: PC1 e PC2:

$$\Delta = q_{20; 63; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)} = 5,24 \sqrt{\frac{0,3131}{2}} = 2,0733$$

PC1	85,3917	a
PC2	85,0167	a

Médias seguidas pela mesma letra não diferem, ao nível de 5% de significância, pelo teste de Tukey.

b) Na comparação entre dois tratamentos regulares de um mesmo experimento

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2}{J} \hat{\sigma}^2 = \frac{2}{4} \times 1,8787 = 0,9394$$

Teste de Tukey para comparar as procedências regulares pertencentes ao experimento 1.

$$\Delta = q_{20; 63; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)} = 5,24 \sqrt{\frac{0,9394}{2}} = 1,8787 = 3,5912$$

Procedências	Média ajustada	
P6	90,5292	a
P3	89,9292	ab
P1	86,9042	bc
P4	86,4792	bc
P2	84,0042	c
P5	83,4042	c

Médias seguidas pela mesma letra não diferem, ao nível de 5% de significância, pelo teste de Tukey.

c) Na comparação entre dois tratamentos regulares de diferentes experimentos

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2}{J} \left(1 + \frac{1}{n_c}\right) \hat{\sigma}^2 = \frac{2}{4} \left(1 + \frac{1}{2}\right) \times 1,8787 = 1,4090$$

n_c : número de tratamentos comuns

O teste de Tukey para comparar a procedência regular (P6) do experimento 1 com a procedência regular P11 do experimento 2.

$$\Delta = q_{20; 63; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)} = 5,24 \sqrt{\frac{1,4090}{2}} = 4,3982$$

P6	90,5292	a
P11	92,5542	a

Médias seguidas pela mesma letra não diferem, ao nível de 5% de significância, pelo teste de Tukey.

d) Na comparação entre um tratamento comum com um regular

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{1}{J} \left(1 + \frac{1}{n_c} + \frac{1}{n_g} - \frac{1}{n_c n_g} \right) \hat{\sigma}^2 = \frac{1}{4} \left(1 + \frac{1}{2} + \frac{1}{3} - \frac{1}{2 \times 3} \right) \times 1,8787 = 0,7828$$

O teste de Tukey, comparando a procedência comum (PC1) com a regular P11.

$$\Delta = q_{20; 63; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)} = 5,24 \sqrt{\frac{0,7828}{2}} = 3,2782$$

PC1	85,3917 a
P11	92,5542 b

Médias seguidas pela mesma letra não diferem, ao nível de 5% de significância, pelo teste de Tukey.

4.5. COMPONENTES DE VARIÂNCIA PARA O ENSAIO CASUALIZADO EM BLOCOS COM K AMOSTRAS POR PARCELA.

$$x_{ijk} = m + t_i + b_j + e_{ij} + \varepsilon_{ijk}$$

m : média geral;

t_i : efeito de tratamentos, em que $i = 1, \dots, I$ (em que I : número de tratamentos);

b_j : efeito fixo de bloco, em que $j = 1, \dots, J$ (em que J : número de blocos);

e_{ij} : erro experimental (ou erro entre), em que $j = 1, \dots, J$ (em que J : número de parcelas por tratamento), tal que: $E(e_{ij}^2) = \sigma_e^2$;

ε_{ijk} : erro amostral (ou erro dentro), em que $k = 1, \dots, K$ (em que K : número de amostras por parcela), tal que $E(\varepsilon_{ijk}^2) = \sigma_\varepsilon^2$.

Tabela 4.5.1. Componentes de variância para o ensaio casualizado em blocos com K amostras por parcela.

F.V.	G.L.	Q.M.	E(Q.M.)		
			Aleatório	Fixo	Misto
Tratamentos	$I - 1$	V_1	$\sigma_\varepsilon^2 + K\sigma_e^2 + JK\phi_t^2$	$\sigma_\varepsilon^2 + K\sigma_e^2 + JK\phi_t$	$\sigma_\varepsilon^2 + K\sigma_e^2 + JK\phi_t$
Blocos	$J - 1$	V_2	$\sigma_\varepsilon^2 + K\sigma_e^2 + IK\sigma_b^2$	$\sigma_\varepsilon^2 + K\sigma_e^2 + IK\phi_b$	$\sigma_\varepsilon^2 + K\sigma_e^2 + IK\phi_b$
Resíduo	$(I-1)(J-1)$	V_3	$\sigma_\varepsilon^2 + K\sigma_e^2$	$\sigma_\varepsilon^2 + K\sigma_e^2$	$\sigma_\varepsilon^2 + K\sigma_e^2$
Parcelas	$IJ - 1$				
Entre amostras d. parcelas	$IK(K-1)$	V_4	σ_ε^2	σ_ε^2	σ_ε^2
Total	$IJK-1$				

$$\phi_t = \frac{\sum_{i=1}^I t_i^2}{I} \quad \phi_b = \frac{\sum_{j=1}^J b_j^2}{J}$$

4.6. ANÁLISE DE EXPERIMENTOS DE ADUBAÇÃO

A análise de variância pressupõe a independência dos diversos tratamentos. Por outro lado, quando esta condicionante não se verifica, então a análise de variância pode não ser válida, haja vista que ela deve refletir a dependência entre os tratamentos.

Quando os tratamentos são quantitativos, tais como — como datas de semeadura, doses crescentes de fertilizantes e inseticidas, entre outros —, haverá uma correspondência funcional, denominada regressão, que associará os valores dos tratamentos com os dados mensurados da variável resposta.

4.6.1. Análise considerando a regressão polinomial

Exercício 4.6.1.1. Considerar um experimento casualizado em blocos de adubação com P_2O_5 , considerando as doses 0, 30, 60, 90 e 120 kg ha⁻¹.

Tabela 4.6.1.1. Produção de madeira (dados simulados) em m³ ha⁻¹.

Doses	I	II	III	IV	V	Total
0	63	58	69	50	49	289
30	90	105	87	71	78	431
60	111	114	112	115	104	556
90	125	110	120	115	105	575
120	115	113	120	130	121	599
Total	504	500	508	481	457	2.450,0

a) Análise de variância não considerando a regressão

Tabela 4.6.1.2. Análise de variância sem considerar a regressão.

F.V.	G.L.	S.Q.	Q.M.	<i>F</i>
Blocos	4	358,0	89,5	
Doses	4	13.468,8	3.367,2	49,87 **
Resíduo	16	1.099,2	68,7	
Total	24	14.926,0		

$$F_{4;16;0,05} = 3,01$$

Verifica-se haver efeito significativo, considerando as doses aplicadas.

b) Cálculo dos efeitos: linear, quadrático, cúbico e de quarto grau.

A técnica dos polinômios ortogonais consiste em desdobrar os quatro graus de liberdade dos tratamentos (doses) e estimar separadamente os efeitos de primeiro, segundo, terceiro (cúbico) e quarto graus, entre outros.

Considerando que as cinco doses são igualmente espaçadas, os coeficientes dos contrastes podem ser obtidos em tabelas específicas. Nesse exemplo de aplicação, tem-se:

1º grau	2º Grau	3º Grau	4º Grau	Total
-2	+2	-1	+1	289
-1	-1	+2	-4	431
0	-2	0	+6	556
+1	-1	-2	-4	575
+2	+2	+1	+1	599
$k_1 = 10$	$k_2 = 14$	$k_3 = 10$	$k_4 = 10$	

A fim de calcular a soma dos quadrados de cada componente, existem as seguintes fórmulas da teoria dos polinômios ortogonais:

$$SQ_{LINEAR} = \frac{(\sum_{i=1}^p C_i T_i)^2}{JK_1} = \frac{(-2 \times 289 - 1 \times 431 + 0 \times 556 + 1 \times 575 + 2 \times 599)^2}{5 \times 10} = \frac{(764)^2}{5 \times 10} = 11.673,92$$

$$SQ_{QUADRÁTICO} = \frac{(\sum_{i=1}^p C_i T_i)^2}{JK_2} = \frac{(2 \times 289 - 1 \times 431 - 2 \times 556 - 1 \times 575 + 2 \times 599)^2}{5 \times 14} = \frac{(-342)^2}{5 \times 14} = 1.670,9143$$

$$SQ_{CÚBICO} = \frac{(\sum_{i=1}^p C_i T_i)^2}{JK_3} = \frac{(-1 \times 289 + 2 \times 431 + 0 \times 556 - 2 \times 575 + 1 \times 599)^2}{5 \times 10} = \frac{(22)^2}{5 \times 10} = 9,68$$

$$SQ_{4^o \text{ GRAU}} = \frac{(\sum_{i=1}^p C_i T_i)^2}{JK_4} = \frac{(1 \times 289 - 4 \times 431 + 6 \times 556 - 4 \times 575 + 1 \times 599)^2}{5 \times 70} = \frac{(200)^2}{5 \times 70} = 114,2857$$

Análise de variância, considerando os quatro efeitos

Tabela 4.6.1.3. ANOVA considerando os quatro efeitos.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	4	358,0000	89,5000	
Doses	(4)	(13.468,8000)	3.367,2000	49,01**
Efeito linear	1	11.673,9200	11.673,9200	169,93**
Efeito quadrático	1	1.670,9143	1.670,9143	24,32 **
Efeito cúbico	1	9,6800	9,6800	0,14 <i>ns</i>
Efeito de 4º grau	1	114,2857	114,2857	2,16 <i>ns</i>
Resíduo	16	1.099,2000	68,7000	1,66 <i>ns</i>
Total	24	14.926,0000		

Em que:

p : número de doses;

J : número de blocos;

C_i : representa os coeficientes dos polinômios ortogonais;

T_i : reporta os totais dos tratamentos nas $p=5$ doses;

$K_1; K_2; K_3; K_4$: são as somas de quadrados dos coeficientes dos polinômios ortogonais.

$$SQ_{DOSES} = 11.673,92 + 1.670,9143 + 9,68 + 114,2857 = 13.468,8$$

$$F_{1;16;0,05} = 4,49 \quad F_{1;16;0,01} = 8,53$$

Os resultados do teste F mostram que apenas os componentes de primeiro grau e de segundo grau apresentaram significância. Logo, deverá ser ajustada a regressão correspondente ao maior grau significativo, ou seja, o modelo do trinômio de segundo grau: $Y_i = a + bX + cX^2 + e_i$. Em que:

$$\hat{Y}_i = 57,66857 + 1,16076X - 0,00543X^2$$

Tabela 4.6.1.4. Estimativa dos parâmetros.

Variáveis	G.L	Parâmetros Estimados	Erro padrão	Valor de “t”	Pr > t
Intercepto	1	57,66857	3,56812	16,16	< 0,0001
Doses	1	1,16076	0,14089	8,24	< 0,0001
DOSES2	1	-0,00543	0,00113	-4,82	< 0,0001

Por outro lado, considerando que $\hat{Y}_i = 57,66857 + 1,16076X - 0,00543X^2$ é a expressão analítica que estima a produção da madeira, em função da dosagem do fertilizante, então é recomendável estimar o valor da dose que promove a máxima produção. No exemplo apresentado, tem-se:

$$f'(X) = 1,16076 - 2 \times 0,00543X^* = 0$$

$$X^* = 106,88 \text{ kg ha}^{-1}$$

Conclui-se que a quantidade recomendada de P_2O_5 para se obter a máxima produção é de 106,88 kg ha⁻¹.

Entretanto, o procedimento ideal para se obter a dose recomendável deveria considerar o fator econômico, ou seja, maximizar a receita líquida (RL), a qual é dada pela expressão:

$$RL = w\hat{Y} - df - cfX$$

Na qual:

w : receita obtida pela venda da madeira por (m³ ha⁻¹);

df : despesas fixas;

cf : custo do fertilizante P_2O_5 ,

Derivando-se a função RL em relação a X , tem-se:

$$RL' = w\hat{Y}' - cf$$

No caso da função quadrática: $Y = a + bX + cX^2$, em que a primeira derivada é:

$$Y' = b + 2cX :$$

$$RL' = w(b + 2cX) - cf$$

$$w(b + 2cX^*) - cf = 0 \rightarrow X^* = \frac{bw - cf}{-2cw}$$

A expressão de X^* em função da razão R , ($R = \frac{w}{cf}$), entre o valor da receita obtido pela venda da madeira (w) e o custo da adubação (cf), resulta em:

$$X^* = \frac{bRcf - cf}{-2cRcf} \quad X^* = \frac{bR - 1}{-2cR}$$

Adotando-se a razão $R = 5$ (Do exercício 4.6.1.1), em que:

$$\hat{Y}_i = 57,66857 + 1,16076X - 0,00543X^2$$

$$X^* = \frac{bR - 1}{-2cR} = \frac{1,16076 \times 5 - 1}{-2 \times 5 \times -0,00543} = 88,4678 \text{ kg ha}^{-1}$$

Por outro lado, a expressão de X^* em função da razão $R^* = \frac{cf}{w}$, isto é, entre o custo da adubação (cf) e o valor da receita obtido pela venda da madeira (w), em^w que:

$R^* = \frac{1}{R}$, resulta em:

$$X^* = \frac{b - R^*}{-2c}$$

Considerando $R^* = 0,20$, assim como $\hat{Y}_i = 57,66857 + 1,16076X - 0,00543X^2$, tem-se:

$$X^* = \frac{b - R^*}{-2c} = \frac{1,16076 - 0,20}{-2 \times -0,00543} = 88,4678 \text{ kg ha}^{-1}$$

De acordo com Pimentel Gomes e Nogueira (1980), uma desvantagem do uso do trinômio do segundo grau é que ele não dá uma boa extrapolação, pois a curva decresce muito rapidamente após o ponto de máximo, o que não ocorre na prática em experimentos com adubação. Assim sendo, na análise de experimentos de adubação, a equação de Mitscherlich é muito mais utilizada.

4.6.2. Análise considerando a equação de Mitscherlich

$$Y = \beta_0 [1 - 10^{-\beta_1 (X + \beta_2)}] + e_i$$

Em que:

β_0 : Corresponde à produção máxima teórica possível quando a dose do nutriente é aumentada indefinidamente;

β_1 : É um parâmetro típico do nutriente. É denominado de coeficiente de eficácia e importante para comparar tipos de nutrientes;

β_2 : É o teor do nutriente assimilável pela planta contido no solo.

Essa função é sempre crescente, pois sua derivada, dada pela fórmula 4.6.2.1, é sempre positiva, denotando que não existe máximo.

$$Y' = \ln 10 \times \beta_0 \beta_1 [10^{-\beta_1 (X + \beta_2)}] \quad 4.6.2.1.$$

Por outro lado, do ponto de vista da análise de experimentos de adubação, muitas vezes não é suficiente apenas calcular o máximo de uma função. Fundamental é levar em consideração o fator econômico. Dessa maneira, é importante se obter a produção que maximiza a renda líquida.

Considerando w o preço da madeira por ($\text{m}^3 \text{ha}^{-1}$) e Y o volume ($\text{m}^3 \text{ha}^{-1}$), wY representará a receita bruta. Considerando as despesas, compostas de uma parte fixa, para a execução do experimento (df) e o custo na compra do fertilizante (cf), o qual será aplicado de acordo com a dosagem (X), a fórmula da receita líquida será dada pela função Z , exposta a seguir:

$$Z = wY - df - cfX$$

Em que:

w : receita obtida pela venda da madeira por ($\text{m}^3 \text{ha}^{-1}$);

Y : volume de madeira ($\text{m}^3 \text{ha}^{-1}$) estimado pela função de crescimento, a partir da equação de Mitscherlich;

df : despesas fixas;

cf : custo do fertilizante;

X : dosagens de adubo aplicadas.

Tomando-se a primeira derivada da função Z em relação à variável tempo X , tem-se:

$$Z' = \frac{dZ}{dX} = wY' - cf$$

Da definição de derivada, para a função $y = ck^u$, tal que $u = f(X)$, logo:

$$y' = ck^u \times \ln k \times u'$$

Dado que $Y = \beta_0[1 - 10^{-\beta_1(X+\beta_2)}] = \beta_0 - \beta_0 10^{-\beta_1(X+\beta_2)}$, tem-se:

$$\frac{dY}{dX} = Y' = 0 - \beta_0 10^{-\beta_1(X+\beta_2)} \times \ln 10 \times -\beta_1$$

$$Y' = \ln 10 \times \beta_0 \beta_1 [10^{-\beta_1(X+\beta_2)}]$$

Então:

$$Z' = 2,302585 w[\beta_0 \beta_1 10^{-\beta_1(X+\beta_2)}] - cf$$

Anulando a primeira derivada $Z' = 0$, obtém-se:

$$Z' = 2,302585 w[\hat{\beta}_0 \hat{\beta}_1 10^{-\hat{\beta}_1(X^*+\hat{\beta}_2)}] - cf = 0$$

Em que:

X^* : é a dosagem de adubo economicamente aconselhável. Logo,

$$2,302585 w \hat{\beta}_0 \hat{\beta}_1 10^{-\hat{\beta}_1(X^*+\hat{\beta}_2)} = cf$$

$$10^{-\hat{\beta}_1(X^*+\hat{\beta}_2)} = \frac{cf}{2,302585 \times w \times \hat{\beta}_0 \times \hat{\beta}_1}$$

Tomando-se o logaritmo na base dez de ambos os membros, tem-se:

$$-\hat{\beta}_1(X^* + \hat{\beta}_2) = \log_{10}\left(\frac{cf}{2,302585 \times w \times \hat{\beta}_0 \times \hat{\beta}_1}\right)$$

$$-\hat{\beta}_1 X^* - \hat{\beta}_1 \hat{\beta}_2 = \log_{10}\left(\frac{cf}{2,302585 \times w \times \hat{\beta}_0 \times \hat{\beta}_1}\right)$$

$$-\hat{\beta}_1 X^* = \log_{10}\left(\frac{cf}{2,302585 \times w \times \hat{\beta}_0 \times \hat{\beta}_1}\right) + \hat{\beta}_1 \hat{\beta}_2$$

$$\hat{\beta}_1 X^* = -\log_{10}\left(\frac{cf}{2,302585 \times w \times \hat{\beta}_0 \times \hat{\beta}_1}\right) - \hat{\beta}_1 \hat{\beta}_2$$

$$X^* = -\frac{1}{\hat{\beta}_1}[\log_{10}\left(\frac{cf}{2,302585 \times w \times \hat{\beta}_0 \times \hat{\beta}_1}\right)] - \hat{\beta}_2$$

Definindo R como uma razão de custo entre a receita obtida pela venda da madeira (w) e o custo do fertilizante aplicado por hectare (cf), obter-se-á:

$$R = \frac{w}{cf} \rightarrow X^* = -\frac{1}{\hat{\beta}_1} [\log_{10}(\frac{1}{2,302585 \times R \times \hat{\beta}_0 \times \hat{\beta}_1})] - \hat{\beta}_2$$

E, considerando a seguinte propriedade dos logaritmos:

$$\log_a \left(\frac{b}{c} \right) = \log_a b - \log_a c$$

Tem-se:

$$\log_{10} \left(\frac{1}{2,302585 \times R \times \hat{\beta}_0 \times \hat{\beta}_1} \right) = \log_{10} 1 - \log(2,302585 \times R \times \hat{\beta}_0 \times \hat{\beta}_1)$$

$$\log_{10} \left(\frac{1}{2,302585 \times R \times \hat{\beta}_0 \times \hat{\beta}_1} \right) = -\log(2,302585 \times R \times \hat{\beta}_0 \times \hat{\beta}_1)$$

Logo:

$$X^* = \frac{1}{\hat{\beta}_1} \log_{10}(2,302585 R \hat{\beta}_0 \hat{\beta}_1) - \hat{\beta}_2 \quad 4.6.2.2.$$

A expressão 4.6.2.2, obtida a partir da equação de Mitscherlich, estima a dose de adubo economicamente aconselhável em função da razão de custo R , definida como a razão de custo entre a receita obtida pela venda da madeira (w) e o custo do fertilizante por hectare (cf).

A expressão 4.6.2.2, em função da razão entre o custo da adubação (cf) e o valor da receita obtido pela venda da madeira (w), ou seja, $R^* = \frac{cf}{w}$, resulta:

$$X^* = \frac{1}{\hat{\beta}_1} \log_{10} \left(2,302585 \frac{1}{R^*} \hat{\beta}_0 \hat{\beta}_1 \right) - \hat{\beta}_2$$

Pimentel Gomes e Nogueira (1980) apresentam a seguinte fórmula:

$$X^* = \frac{1}{\hat{\beta}_1} \log_{10} \left(\frac{w \hat{\beta}_0 \hat{\beta}_1}{cf \log e} \right) - \hat{\beta}_2$$

Considerando $R = \frac{w}{cf}$, então:

$$X^* = \frac{1}{\hat{\beta}_1} \log_{10} \left(\frac{R \hat{\beta}_0 \hat{\beta}_1}{0,43429448} \right) - \hat{\beta}_2$$

Exercício 4.6.2.1. Considerando um experimento de adubação com P_2O_5 , com a aplicação das doses 0, 30, 60, 90 e 120 kg ha⁻¹, casualizado em blocos. A partir dos dados simulados do exercício 4.6.1.1, obter a dose economicamente aconselhável usando a equação de Mitscherlich.

A partir dos dados da tabela 4.6.1.1, de acordo com a análise feita pelo programa SAS, **Prog17**, em anexo no apêndice B, tem-se:

$$Y_i = \beta_0 [1 - 10^{-\beta_1(X+\beta_2)}] + e_i \rightarrow \hat{Y}_i = 126,1[1 - 10^{-0,00923(X+28,3347)}]$$

Em que:

$\hat{\beta}_0 = 126,1$: significa a produção máxima teórica possível quando a dose do nutriente for aumentada indefinidamente;

$\hat{\beta}_1 = 0,00923$: é o valor do coeficiente de eficácia do nutriente testado;

$\hat{\beta}_2 = 28,3347$: é o teor do nutriente contido no solo assimilável pela espécie.

Tabela 4.6.2.1. Análise de regressão no SAS.

Iteração	b0	b1	b2	Soma de Quadrados
0	124,0	0,0100	28,0000	1664,1
1	126,1	0,00910	28,6323	1624,6
2	126,0	0,00924	28,2939	1622,5
3	126,1	0,00923	28,3386	1622,5
4	126,1	0,00923	28,3343	1622,5
5	126,1	0,00923	28,3347	1622,5

Considerando que $R = \frac{w}{cf} = 5$, significando que o custo da adubação corresponderá a 20 % da receita obtida pela venda da madeira, tem-se:

$$X^* = \frac{1}{\hat{\beta}_1} \log_{10} (2,302585 R \hat{\beta}_0 \hat{\beta}_1) - \hat{\beta}_2$$

$$X^* = \frac{1}{0,00923} \log_{10} (2,302585 \times 5 \times 126,1 \times 0,00923) - 28,3349$$

$$X^* = 93,7780 \text{ kg ha}^{-1}$$

Considerando a alternativa apresentada por $R^* = \frac{cf}{w} = \frac{1}{5} = 0,20$, ter-se-á:

$$X^* = \frac{1}{\hat{\beta}_1} \log_{10} (2,302585 \frac{1}{R^*} \hat{\beta}_0 \hat{\beta}_1) - \hat{\beta}_2$$

$$X^* = \frac{1}{0,00923} \log_{10} (2,302585 \times \frac{1}{0,20} \times 126,1 \times 0,00923) - 28,3349 = 93,7780 \text{ kg ha}^{-1}$$

Usando a fórmula proposta por Pimentel Gomes e Nogueira (1980), para $R = \frac{w}{cf} = 5$, ter-se-á:

$$X^* = \frac{1}{\hat{\beta}_1} \log_{10} \left(\frac{R \hat{\beta}_0 \hat{\beta}_1}{\log e} \right) - \hat{\beta}_2$$

$$X^* = \frac{1}{0,00923} \log_{10} \left(\frac{5 \times 126,1 \times 0,00923}{0,43429448} \right) - 28,3349 = 93,7780 \text{ kg ha}^{-1}$$

Utilizando os dados do exercício 4.6.1.1, para $R^* = 0,20$, e considerando o trinômio do segundo grau, a dose economicamente aconselhável seria de 88,4678 kg ha⁻¹. No caso da equação de Mitscherlich, a dose recomendada seria de 93,77809 kg ha⁻¹.

4.7. BIBLIOGRAFIA

BANZATTO, D. A.; KRONKA, N. **Experimentação agrícola**. Jaboticabal: FUNEP, 1989. 247 p.

CAMPOS, H.; PIEDADE, S. M. S. **Planejamento de experimentos**. 1ª ed. Piracicaba: FEALQ, 1992. 185 p.

MARKUS, R. **Elementos de estatística aplicada**. Porto Alegre: UFRGS, 1971. 329 p.

NETER, J.; WASSERMAN, W.; KUTNER, M. H. **Applied statistical models**. Richard D. Irwin, Inc. Boston, 1990, 1181p.

PADOVANI, C R. **Delineamento de experimentos**. São Paulo: Cultura Acadêmica: Universidade Estadual Paulista, Pró-Reitoria de Graduação, 2014. 128 p.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 14ª ed. São Paulo: Nobel, 2000. 475 p.

PIMENTEL GOMES, F.; GARCIA, C. H. **Estatística aplicada a experimentos agrônômicos e florestais**. Piracicaba: FEALQ (Biblioteca de Ciências Agrárias Luiz de Queiroz - Volume 11), 2002. 309 p.

PIMENTEL GOMES, F.; NOGUEIRA, I. R. **Análise matemática**. 2ª ed. Piracicaba: ESALQ-USP, 1980. 371 p.

SILVA, I. P.; SILVA, J. A. A. **Métodos estatísticos aplicados à pesquisa científica: uma abordagem para profissionais da pesquisa agropecuária**. Universidade Federal Rural de Pernambuco: UFRPE, 2002. 324 p.

STEEL, R. G. D.; TORRIE, J. H. **Principles and procedures of statistics: with special reference to the biological sciences**. New York: Mc Graw-Hill Book Company, Inc., 1960. 481 p.

TUKEY, J. W. One degree of freedom for non-additivity. **International Biometrics Society**, 5(3): 232–242, 1949a.



CAPÍTULO 5

DELINEAMENTO EM QUADRADO LATINO

Este capítulo apresenta definições, conceitos e aplicações sobre a análise de experimentos em Quadrado Latino. Exibe um exemplo de análise de um ensaio 5×5 . Mostra, ainda, uma análise com a aplicação dos testes de Tukey e de Dunnett num ensaio 5×5 com duas testemunhas.

5.1. INTRODUÇÃO

O delineamento em quadrado latino foi idealizado por Sir Ronald A. Fisher na década de 1920, para controlar duas fontes de variação presentes em unidades experimentais, denominadas de duplo bloqueamento. Como o delineamento casualizado em blocos, o quadrado latino é também apropriado para ambientes experimentais heterogêneos.

O delineamento em quadrado latino considera os três princípios básicos da experimentação: repetição, casualização e controle local. Entretanto, o controle local é efetuado visando controlar duas fontes de variação: as linhas e as colunas.

Nos delineamentos em quadrado latino, os tratamentos são distribuídos aleatoriamente nos blocos (linhas e colunas), tal que cada um apareça apenas uma só vez em cada linha e em cada coluna.

No quadrado latino, a heterogeneidade está associada às variações decorrentes de fatores que requerem controle local e, para garantir a homogeneidade dentro dos blocos, necessitam de dois tipos de estratificação de parcelas, como no caso de terrenos com declividades e gradientes de fertilidade em dois sentidos.

A concepção estrutural experimental do quadrado latino estabelece o controle local com dois tipos de blocos completos, ou seja, as linhas (homogêneas para um fator de variabilidade nas unidades experimentais) e as colunas (homogêneas para o outro fator de variação nas parcelas), originando a denominação de sistema duplo de blocos, em que cada bloco, representado por linhas e colunas, terá o mesmo número de parcelas quantos forem os tratamentos.

O número de parcelas que constituem o experimento é sempre um quadrado perfeito, $I^2 = 9, 16, 25, 36, 49$ etc, haja vista que cada linha e cada coluna (blocos completos) deverão conter todos os tratamentos uma única vez. Essa exigência faz com que o número de tratamentos (I) seja igual ao número de repetições e, conseqüentemente, aos números de linhas e de colunas.

O delineamento em quadrado latino possui uma aplicação muito restrita, pois apresenta pouca flexibilidade quando o número de tratamentos é pequeno ou alto. Por exemplo, um ensaio com menos de cinco tratamentos apresenta um número pequeno de graus de liberdade para o resíduo, resultando em baixa estimativa de precisão para o erro experimental, diminuindo a sensibilidade dos testes para as comparações de médias de tratamentos.

Por outro lado, se o número de tratamentos for superior a oito, haverá um aumento considerável do número de repetições e de parcelas do experimento, o que poderá impossibilitar a instalação no campo.

Contudo, para contornar esse problema, podem-se utilizar dois ou mais ensaios em quadrado latino e assim obter, por meio da análise conjunta, a estimativa da variância aleatória (erro experimental), haja vista o aumento do número de graus de liberdade, promovendo o aumento da sensibilidade dos testes a serem aplicados.

A perda de parcelas nos experimentos em quadrado latino poderá ocasionar dificuldades e complicar a análise estatística, pois a eliminação de tratamento(s), linha(s) e/ou coluna(s), pode promover a descaracterização conceitual do delineamento.

Por outro lado, de acordo com Duarte (1996), é possível, conforme o tipo de perda, transformar a ANOVA para casualizado em blocos, desprezando-se o controle feito pelas linhas ou pelas colunas. Em alguns casos, de acordo com Pimentel Gomes (2000) e Banzatto e Kronka (1989), a solução pode ser obtida pela estimação de observações para as parcelas perdidas, seguida de certos ajustamentos.

Em condições ambientais homogêneas, utilizando-se o delineamento inteiramente casualizado, espera-se uma maior precisão. Por outro lado, quando existirem condições heterogêneas, a precisão será maior utilizando-se o delineamento casualizado em blocos, obviamente no caso de ocorrer um único tipo de estratificação. Contudo, para controlar a variabilidade, caso ocorra a necessidade de um bloqueamento duplo, então deverá ser adotado o quadrado latino.

5.2. DISTRIBUIÇÃO DOS TRATAMENTOS NAS UNIDADES EXPERIMENTAIS

Duarte (1996) comenta que uma forma de casualização que pode contribuir para melhorar a imparcialidade é a da identificação prévia, mediante sorteio, dos tratamentos identificados por letras ou números. Esse é um procedimento importante no planejamento de séries de experimentos em quadrado latino, em que se recomenda um sorteio para cada experimento.

Na distribuição dos tratamentos, ocorre que cada um deles deve aparecer uma única vez em cada linha e em cada coluna, caracterizando a propriedade denominada de restrição dupla na casualização. Tal disposição das unidades nas linhas e colunas visa garantir uma distribuição sistemática dos tratamentos às unidades experimentais. Exemplificando essa alocação, eis o exemplo de um experimento com 6 tratamentos (A, B, C, D, E, F) delineado em quadrado latino, em que serão necessárias 36 parcelas. O quadro 5.2.1 mostra a distribuição inicial.

Quadro 5.2.1. Distribuição sistemática dos tratamentos em um quadrado latino 6×6.

	Coluna 1	Coluna 2	Coluna 3	Coluna 4	Coluna 5	Coluna 6
Linha 1	A	B	C	D	E	F
Linha 2	F	A	B	C	D	E
Linha 3	E	F	A	B	C	D
Linha 4	D	E	F	A	B	C
Linha 5	C	D	E	F	A	B
Linha 6	B	C	D	E	F	A

Para assegurar o princípio da casualização, ou imparcialidade na alocação dos tratamentos às parcelas (independência dos efeitos aleatórios sobre os tratamentos), processa-se uma redistribuição aleatória das linhas e colunas. Desse modo, sorteando a ordem das linhas, a sequência prévia fornecida pela distribuição sistemática é alterada.

O quadro 5.2.2 mostra a distribuição do quadrado latino 6×6 após o sorteio de linhas, assim como o quadro 5.2.3 apresenta a distribuição após o sorteio de colunas.

Quadro 5.2.2. Distribuição após o sorteio de linhas em um quadrado latino 6×6.

	Coluna 1	Coluna 2	Coluna 3	Coluna 4	Coluna 5	Coluna 6
Linha 6	B	C	D	E	F	A
Linha 4	D	E	F	A	B	C
Linha 3	E	F	A	B	C	D
Linha 5	C	D	E	F	A	B
Linha 2	F	A	B	C	D	E
Linha 1	A	B	C	D	E	F

Quadro 5.2.3. Distribuição após o sorteio de colunas em um quadrado latino 6×6.

	Coluna 4	Coluna 5	Coluna 3	Coluna 2	Coluna 6	Coluna 1
Linha 6	E	F	D	C	A	B
Linha 4	A	B	F	E	C	D
Linha 3	B	C	A	F	D	E
Linha 5	F	A	E	D	B	C
Linha 2	C	D	B	A	E	F
Linha 1	D	E	C	B	F	A

O último quadro corresponde à distribuição definitiva dos tratamentos às parcelas em um quadrado latino 6×6.

5.3. ANÁLISE DE VARIÂNCIA

No quadrado latino, a análise de variância é denominada de análise segundo três critérios de classificação, isto é, tratamentos, linhas e colunas, pressupondo que a variabilidade total dos dados observados, para a variável resposta, nas unidades experimentais, pode ser decomposta nos efeitos decorrentes de causas controladas, ou seja, tratamentos, linhas, colunas e de causas não controladas (erro ou resíduo), conforme o modelo matemático:

$$x_{ijk} = m + t_i + l_j + c_k + e_{ijk}$$

Em que:

x_{ijk} : observação na unidade experimental que recebeu o i -ésimo tratamento, $i = 1, 2, \dots, I$, na j -ésima linha, $j = 1, 2, \dots, J$ e na k -ésima coluna, $k = 1, 2, \dots, K$.

m : média geral;

t_i : efeito do i -ésimo tratamento;

l_j : efeito da j -ésima linha;

c_k : efeito da k -ésima coluna;

e_{ijk} : erro na unidade experimental.

Exercício 5.3.1. Considerar um estudo de competição de cinco espécies florestais, num ensaio em um Quadrado Latino de 5×5, instalado em um terreno apresentando inclinação no sentido das colunas e variação de fertilidade no sentido das linhas. No quadro abaixo, estão as produções de madeira no final da rotação (dados simulados).

Tabela 5.3.1. Produção de madeira em m³ por parcela.

	C1	C2	C3	C4	C5	Total
L1	E2=30,1	E5=59,6	E4=35,5	E3=60,0	E1=45,9	231,1
L2	E3=48,5	E1=63,6	E5=38,1	E4=46,0	E2=49,5	245,7
L3	E1=47,2	E4=46,5	E3=58,1	E2=31,4	E5=32,1	215,3
L4	E4=41,2	E2=45,7	E1=48,5	E5=49,5	E3=46,8	231,7
L5	E5=31,5	E3=82,4	E2=46,7	E1=58,9	E4=40,6	260,1
Total	198,5	297,8	226,9	245,8	214,9	1.183,9

n : número de tratamentos, de linhas e de colunas;

T_i : Total do i -ésimo tratamento;

L_j : Total da j -ésima linha;

C_k : Total da k -ésima coluna.

a) Obtenção das somas de quadrados

Soma de quadrados total:

$$SQ_{TOTAL} = \sum_{i,j,k}^n x_{ijk}^2 - C = (30,1^2 + 59,6^2 + \dots + 58,9^2 + 40,6^2) - C = 3.416,7416$$

$$\sum_{i,j,k}^n x_{ijk}^2 = 59.481,51 \quad C = \frac{(\sum_{i,j,k}^n x_{ijk})^2}{n^2} = \frac{1.183,9^2}{5^2} = 56.064,7684$$

Soma de quadrados das linhas:

$$SQ_{LINHAS} = \frac{\sum_{j=1}^n L_j^2}{n} - C = \frac{231,1^2 + 245,7^2 + 215,3^2 + 231,7^2 + 260,1^2}{5} - C = 228,5696$$

Soma de quadrados das colunas:

$$SQ_{COLUNAS} = \frac{\sum_{k=1}^n C_k^2}{n} - C = \frac{198,5^2 + 297,8^2 + 226,9^2 + 245,8^2 + 214,9^2}{5} - C = 1.169,3016$$

Soma de quadrados dos tratamentos:

	E1	E2	E3	E4	E5
1	45,9	30,1	60,0	35,5	59,6
2	63,6	49,5	48,5	46,0	38,1
3	47,2	31,4	58,1	46,5	32,1
4	48,5	45,7	46,8	41,2	49,5
5	58,9	46,7	82,4	40,6	31,5
Total	264,1	203,4	295,8	209,8	210,8

$$SQ_{TRAT} = \frac{\sum_{i=1}^n T_i^2}{n} - C = \frac{264,1^2 + 203,4^2 + 295,8^2 + 209,8^2 + 210,8^2}{5} - C = 1.349,3696$$

b) Análise de variância

Tabela 5.3.2. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Linhas	4	228,5696	57,1424	1,02 <i>ns</i>
Colunas	4	1.169,3016	292,3254	5,24 *
Tratamentos	4	1.349,3696	337,3424	6,05**
Resíduos	12	669,5008	55,7917	
Total	24	3.416,7416		

$$F_{4;12;0,05} = 3,26 \quad F_{4;12;0,01} = 5,41$$

Os resultados da análise de variância, considerando o teste F , mostram que há argumento para rejeitar a igualdade de médias entre os tratamentos. Por outro lado, existe significância, referente às colunas, e não significância, concernente às linhas.

Tabela 5.3.3. ANOVA realizada pelo SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Trat.	4	1.349.369600	337.342400	6.05	0.0067
Linhas	4	228.569600	57.142400	1.02	0.4338
Colunas	4	1.169,301600	292,325400	5,24	0,0112
Resíduo	12	669,500800	55,791733		
Total	24	3.416,741600			

c) Teste de comparações múltiplas de Tukey:

$$\Delta = q_{5,12;0,05} \times \sqrt{\frac{QM_{RESÍDUO}}{n}} = 4,51 \times \frac{\sqrt{55,7917}}{\sqrt{5}} = 15,0652$$

Teste de Tukey	
Espécie	Média
E3	59,16 a
E1	52,82 ab
E5	42,16 b
E4	41,96 b
E2	40,68 b

Médias seguidas pela mesma letra, para um nível de significância de $\alpha = 0,05$, não diferem pelo teste de Tukey. As espécies E3 e E1 apresentaram resultados não significativos. É recomendável usar a espécie E3.

Alfa (α)	0,05
Graus de liberdade do resíduo	12
Quadrado médio do resíduo	55,79173
Valor crítico da amplitude total estudentizada	4,50760
Diferença mínima de significância	15,057

Tukey (Agrupamento)		Média	N	Trat.
	A	59,160	5	E3
B	A	52,820	5	E1
B		42,160	5	E5
B		41,960	5	E4
B		40,680	5	E2

As análises foram feitas pelo programa SAS, **Prog18**, em anexo no apêndice B.

Exercício 5.3.2. Considerar um estudo de adubação em quadrado latino 5×5 , instalado em um terreno apresentando inclinação no sentido das colunas e variação de fertilidade no sentido das linhas, visando estimar o crescimento de uma determinada espécie florestal. Considerar, ainda, A e B como testemunhas e C, D, E como tratamentos comuns. No quadro abaixo, constam as produções de madeira em m^3 por parcela (dados simulados) e as medidas no final da rotação economicamente aconselhável.

Tabela 5.3.4. Dados do volume em $m^3 ha^{-1}$.

	C1	C2	C3	C4	C5	Total
L1	B=100,0	A=79,1	C=91,6	D=118,0	E=105,5	494,2
L2	A=98,5	B=98,5	D=115,2	E=119,4	C=94,4	526,0
L3	E=113,8	C=91,6	A=115,2	B=98,5	D=94,4	513,5
L4	C=74,9	D=90,2	E=104,0	A=84,7	B=91,6	445,4
L5	D=116,6	E=111,1	B=112,4	C=115,2	A=97,1	552,4
Total	503,8	470,5	538,4	535,8	483,0	2.531,5

A: Sem adubação

B: Com adubação mineral

C: Adubação orgânica1

D: Adubação orgânica2

E: Adubação orgânica3

$$SQ_{TOTAL} = 100,0^2 + 79,1^2 + \dots + 115,2^2 + 97,1^2 - \frac{2.531,5^2}{25}$$

$$SQ_{TOTAL} = 260.169,21 - \frac{2.531,5^2}{25} = 3.829,52$$

$$SQ_{LINHAS} = \frac{494,2^2 + 526,6^2 + 513,5^2 + 445,4^2 + 552,4^2}{5} - \frac{2.531,5^2}{25}$$

$$SQ_{LINHAS} = \frac{1.288.118,81}{5} - \frac{2.531,5^2}{25} = 1.284,0720$$

$$SQ_{COLUNAS} = \frac{503,8^2 + 470,5^2 + 538,4^2 + 535,8^2 + 483,0^2}{5} - \frac{2.531,5^2}{25}$$

$$SQ_{COLUNAS} = \frac{1.285.429,89}{5} - \frac{2.531,5^2}{25} = 746,2880$$

A	B	C	D	E
79,1	100,0	91,6	118,0	105,5
98,5	98,5	94,4	115,2	119,4
115,2	98,5	91,6	94,4	113,8
84,7	91,6	74,9	90,2	104,0
97,1	112,4	115,2	116,6	111,1
474,6	501,0	467,7	534,4	553,8

$$SQ_{TRAT} = \frac{474,6^2 + 501,0^2 + 467,7^2 + 534,4^2 + 553,8^2}{5} - \frac{2.531,5^2}{25}$$

$$SQ_{TRAT} = \frac{1.287.267,25}{5} - \frac{2.531,5^2}{25} = 1.113,76$$

$$SQ_{RESÍDUO} = SQ_{TOTAL} - SQ_{TRAT} - SQ_{LINHAS} - SQ_{COLUNAS} = 685,4$$

$$SQ_{TRAT.COMUNS} = \frac{467,7^2 + 534,4^2 + 553,8^2}{5} - \frac{1.555,9^2}{15} = 815,8973$$

$$SQ_{TESTEMUNHAS} = \frac{474,6^2 + 501,0^2}{5} - \frac{975,6^2}{10} = 69,6960$$

$$SQ_{COMUNS VS TESTEMUNHAS} = \frac{1.555,9^2}{15} + \frac{975,6^2}{10} - \frac{2.531,5^2}{25} = 228,1667$$

Tabela 5.3.5. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Tratamentos	(4)	(1.113,76)	-----	-----
Trat. Comuns	2	815,8973	407,9487	7,14**
Testemunhas A e B	1	69,6960	69,6960	1,22 <i>ns</i>
Comuns vs Test.	1	228,1667	228,1667	3,99 <i>ns</i>
Linhas	4	1.284,0720	321,0180	5,62
Colunas	4	746,2880	186,5720	3,27
Resíduo	12	685,4000	57,1167	
Total	24	3829,5200		

$$F_{2; 12; 0,05} = 3,89 \quad F_{1; 12; 0,05} = 4,75 \quad F_{2; 12; 0,01} = 6,93$$

Os resultados da análise de variância, considerando o teste F , mostram que há argumento para não rejeitar a igualdade de médias entre as duas testemunhas (A e B). Contudo, eles demonstram que existe diferença significativa na comparação entre as médias dos tratamentos C, D e E.

a) Teste de Tukey para as comparações das médias dos tratamentos comuns C, D e E.

$$\Delta = q_{3; 12; 0,05} \sqrt{\frac{QM_{RES}}{J}} = 3,77 \sqrt{\frac{57,1167}{5}} = 12,74$$

$$\hat{m}_C = \frac{467,7}{5} = 93,54 \quad \hat{m}_D = \frac{534,4}{5} = 106,88 \quad \hat{m}_E = \frac{553,8}{5} = 110,76$$

Trat. Comuns	Médias	
E	110,76	a
D	106,88	a
C	93,54	b

Médias seguidas pela mesma letra, para o nível de significância de $\alpha = 0,05$, não diferem pelo teste de Tukey. Os tratamentos E e D apresentaram os melhores resultados.

b) Comparações dos tratamentos comuns C, D e E versus as testemunhas A e B, pelo teste de Dunnett.

$$\hat{m}_A = \frac{470,5}{5} = 94,1 \quad \hat{m}_B = \frac{503,8}{5} = 100,76$$

$$\Pr[\hat{Y} \pm t_{(I-1, n-I, \alpha)} \sqrt{\hat{V}(\hat{Y})}] = 1 - \alpha$$

$I-1$: número de tratamentos menos a testemunha;

$$\Delta = t_{3; 12; 0,05} \sqrt{\frac{2 \times QM_{RES}}{5}} = 2,68 \sqrt{\frac{2 \times 57,1167}{5}} = 12,81$$

b1) Comparações dos tratamentos comuns C, D e E versus a testemunha A.

Trat. Comuns	Médias	Médias-94,1	
E	110,76	16,66	*
D	106,88	12,78	ns
C	93,54	-0,56	ns

Conclui-se que apenas o tratamento E apresentou significância, ao nível de significância $\alpha = 0,05$, em relação à testemunha A (tratamento sem adubação).

b2) Comparações dos tratamentos comuns C, D e E versus a testemunha B (tratamento com adubação mineral).

Trat. Comuns	Médias	Médias-100,76	
E	110,76	10,00	ns
D	106,88	6,12	ns
C	93,54	-7,22	ns

O resultado permite concluir que nenhum tratamento comum apresentou diferença significativa, ao nível de significância $\alpha = 0,05$, em relação à testemunha B (tratamento com adubação mineral).

5.4. BIBLIOGRAFIA

BANZATTO, D. A.; KRONKA, N. **Experimentação agrícola**. Jaboticabal: FUNEP, 1989. 247 p.

DUARTE, J. B. **Princípios sobre delineamentos em experimentação agrícola**. Instituto de Matemática e Física. UFG, 1996. 66 p.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 14^a ed. São Paulo: Nobel, 2000. 475 p.

CAPÍTULO 6

EXPERIMENTOS FATORIAIS

Este capítulo apresenta definições, conceitos e aplicações sobre análise de experimentos fatoriais. Mostra também experimentos fatoriais hierárquicos e a técnica do confundimento. Expõe, ainda, fatoriais fracionários ou fracionados, o planejamento de um fatorial 2^{k-p} e a escolha das combinações a serem testadas. Trata do gerador, da relação definidora e resolução em um experimento fatorial fracionário, além de discorrer acerca de experimentos fatoriais com tratamentos adicionais e exemplos de aplicação.

6.1. CONSIDERAÇÕES GERAIS

A denominação experimentos unifatoriais se aplica ao delineamento de experimentos cujos tratamentos representem variações (ou níveis) de um único fator.

Quando são feitas comparações de tratamentos nos experimentos com um único fator, os demais são considerados constantes. Por exemplo, em um ensaio de competição de espécies florestais, outros fatores — tais como: tipo de solo, características ambientais, tratos silviculturais, adubação, entre outros —, devem ser os mesmos para todas as espécies pesquisadas. Entretanto, existem os experimentos fatoriais que são aplicados em situações em que há necessidade de estudar simultaneamente vários fatores.

Nos ensaios fatoriais, cada subdivisão de um fator é chamada de nível e os tratamentos consistem em todas as combinações entre os diferentes fatores nos seus diversos níveis. Os experimentos fatoriais são mais eficientes do que os experimentos simples. Aqueles permitem tirar conclusões mais amplas no contexto da pesquisa, tornando-a mais rica em termos de informações, haja vista promover a pesquisa de várias possibilidades. O principal problema é que o número de tratamentos aumenta rapidamente com a inclusão de novos fatores, dificultando o uso do delineamento casualizado em blocos e impondo, conseqüentemente, o uso de blocos incompletos e a adesão ao sistema de confundimento. Um fatorial é denominado completo quando todas as combinações estão presentes no experimento e chamado fatorial incompleto ou fracionário quando apenas uma parte das combinações dos fatores é utilizada.

Por exemplo, imaginar um ensaio inteiramente ao acaso para testar os efeitos de três tipos de manejo, considerando duas espécies florestais. No caso, existe um ensaio fatorial 2×3 , em que duas espécies florestais estão combinadas com três diferentes métodos de manejo. O fator espécie possui dois níveis (E1, E2) e o fator manejo possui três níveis (M1, M2, M3), originando as seis combinações (tratamentos):

M1E1	M1E2	M2E1	M2E2	M3E1	M3E2
------	------	------	------	------	------

Caso houvesse um terceiro fator (como adubação) com três níveis (A1, A2, A3), haveria um fatorial $2 \times 3 \times 3$, resultando em 18 combinações (tratamentos) apresentadas a seguir:

M1E1A1	M1E2A1	M2E1A1	M2E2A1	M3E1A1	M3E2A1
M1E1A2	M1E2A2	M2E1A2	M2E2A2	M3E1A2	M3E2A2
M1E1A3	M1E2A3	M2E1A3	M2E2A3	M3E1A3	M3E2A3

Os experimentos fatoriais não configuram um delineamento experimental. Contudo proporcionam esquemas que produzem desdobramentos das somas de quadrados dos tratamentos, visando estudar os componentes envolvidos e suas interações. Eles podem ser instalados em qualquer dos delineamentos experimentais, como, por exemplo, nos casos acima citados: ensaio fatorial 2×3 inteiramente ao acaso ou ensaio fatorial $2 \times 3 \times 3$ casualizado em blocos.

De acordo com o delineamento experimental adotado, nos experimentos fatoriais, após a análise de variância preliminar simples, é feito o desdobramento dos graus de liberdade dos tratamentos, separando os efeitos principais dos fatores e os efeitos das interações entre os fatores estabelecidos. Sendo o exemplo de um fatorial 2×2 ($N \times P$).

	P_0	P_1	Total
N_0	$N_0P_0 = 10$	$N_0P_1 = 20$	30
N_1	$N_1P_0 = 38$	$N_1P_1 = 50$	88
	48	70	118

N_0 : sem nitrogênio N_1 : com nitrogênio

P_0 : sem fósforo P_1 : com fósforo

a) O efeito simples de um fator é a medida da variação que ocorre com a variável resposta em estudo (produção de madeira, por exemplo) correspondente às variações nos níveis desse fator, dentro de cada um dos níveis do outro fator.

a1) Efeito simples de nitrogênio na ausência de fósforo:

$$N d. P_0 = N_1P_0 - N_0P_0 = 38 - 10 = 28$$

a2) Efeito simples de nitrogênio na presença de fósforo:

$$N d. P_1 = N_1P_1 - N_0P_1 = 50 - 20 = 30$$

a3) Efeito simples de fósforo na ausência de nitrogênio:

$$P d. N_0 = N_0P_1 - N_0P_0 = 20 - 10 = 10$$

a4) Efeito simples de fósforo na presença de nitrogênio:

$$P d. N_1 = N_1P_1 - N_1P_0 = 50 - 38 = 12$$

b) Efeito principal de um fator é a medida da variação que ocorre com a variável resposta em estudo (produção de madeira, por exemplo) correspondente às variações nos níveis desse fator, em média de todos os níveis do outro fator. Logo é a média dos efeitos simples desse fator.

b1) Efeito principal de Nitrogênio:

$$N = \frac{N d. P_0 + N d. P_1}{2} = \frac{28 + 30}{2} = 29$$

b2) Efeito principal de Fósforo:

$$P = \frac{P d. N_0 + P d. N_1}{2} = \frac{10 + 12}{2} = 11$$

c) Efeito da interação entre dois fatores é uma medida correspondente às variações nos níveis de um fator, ao passar de um nível a outro do outro fator:

c1) Efeito da interação $N \times P$:

$$N \times P = \frac{N d. P_1 - N d. P_0}{2} = \frac{30 - 28}{2} = 1$$

c2) Efeito da interação $P \times N$:

$$P \times N = \frac{P d. N_1 - P d. N_0}{2} = \frac{12 - 10}{2} = 1$$

Nota-se que $N \times P = P \times N$.

Observando-se o comportamento do nitrogênio na ausência ($N d. P_0$) e também na presença ($N d. P_1$) de fósforo, assim como o do fósforo na ausência ($P d. N_0$) e na presença ($P d. N_1$) de nitrogênio (e analisando os valores do quadro auxiliar), é possível verificar a presença ou ausência da interação. Por outro lado, se o comportamento for o mesmo na ausência e presença, não haverá interação.

Observe-se o exemplo 1, considerando as doses 0 e 20 kg de nitrogênio (N) e fósforo (P) de um experimento de adubação, em que foi medida a altura de uma espécie florestal.

	P_0	P_1	Total
N_0	$N_0 P_0 = 10$	$N_0 P_1 = 15$	25
N_1	$N_1 P_0 = 12$	$N_1 P_1 = 17$	29
	22	32	54

$$N d. P_0 = N_1 P_0 - N_0 P_0 = 12 - 10 = 2$$

$$N d. P_1 = N_1 P_1 - N_0 P_1 = 17 - 15 = 2$$

$$N \times P = \frac{N d. P_1 - N d. P_0}{2} = \frac{2 - 2}{2} = 0$$

Conclui-se não existir interação entre o nitrogênio e o fósforo. A figura 6.1.1 ilustra a inexistência de interação por meio do paralelismo entre as retas.

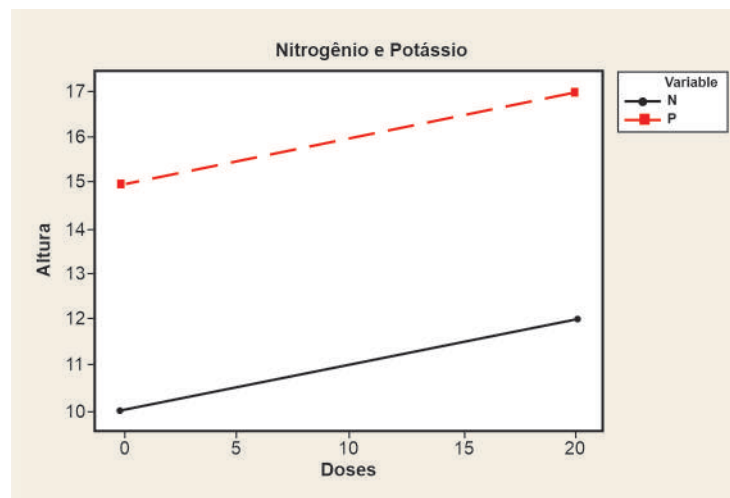


Figura 6.1.1. Paralelismo na tendência de crescimento para N e P .

Observe-se o exemplo2, considerando as doses 0 e 20 kg de nitrogênio (N) e fósforo (P) de um experimento de adubação, em que foi medida a altura de uma espécie florestal.

	P_0	P_1	Total
N_0	$N_0P_0 = 8$	$N_0P_1 = 15$	23
N_1	$N_1P_0 = 10$	$N_1P_1 = 26$	36
	18	41	59

$$N d. P_0 = N_1P_0 - N_0P_0 = 10 - 8 = 2$$

$$N d. P_1 = N_1P_1 - N_0P_1 = 26 - 15 = 11$$

$$N \times P = \frac{N d. P_1 - N d. P_0}{2} = \frac{11 - 2}{2} = 4,5$$

Conclui-se existir interação (positiva) entre o nitrogênio e o fósforo. A figura 6.1.2 ilustra a existência de interação, pois as retas não são paralelas.

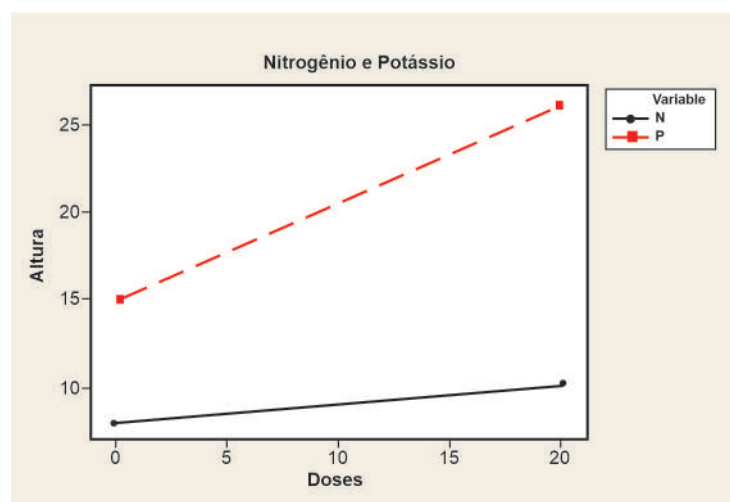


Figura 6.1.2. Não paralelismo na tendência de crescimento para N e P.

As principais vantagens dos arranjos fatoriais em relação aos experimentos simples são as seguintes:

- Possibilita o estudo dos efeitos simples e principais dos fatores, assim como os efeitos das interações entre eles;
- Todas as parcelas são usadas para o cálculo dos efeitos principais dos fatores e dos efeitos das interações.

As principais desvantagens dos experimentos fatoriais em relação aos experimentos simples são as seguintes:

- Os tratamentos são todas as combinações entre os níveis dos diversos fatores, o que torna o número de tratamentos muito alto e, conseqüentemente, dificulta o emprego do delineamento com casualização em blocos. Por outro lado, como há exigência de homogeneidade das unidades experimentais dentro dos blocos, o pesquisador é obrigado a usar a técnica de confundimento;

b) A análise estatística é um pouco mais trabalhosa e a interpretação dos resultados mais complexa, principalmente quando o número de níveis e de fatores é muito grande.

Entretanto existem vários aplicativos computacionais disponíveis que tornam possível efetuar uma boa análise

6.2. ANÁLISE DE VARIÂNCIA DOS EXPERIMENTOS COM DOIS FATORES

A tabela 6.2.1 apresenta o quadro da análise variância para um ensaio inteiramente ao acaso com dois fatores (A e B).

Modelo matemático: $x_{ijk} = m + \alpha_i + \beta_k + (\alpha\beta)_{ik} + e_{ijk}$

J : o número de repetições ($j = 1, \dots, J$);

a : o número de níveis do fator A (α) ($i = 1, \dots, a$);

b : o número de níveis do fator B (β) ($k = 1, \dots, b$).

Tabela 6.2.1. ANOVA para dois fatores A e B (efeitos fixos).

F.V.	G.L.	S.Q.	Q.M.	E(Q.M.)
Fator A	$a-1$	$\frac{1}{bJ} \sum_{i=1}^a A_i^2 - C$	$\frac{SQ_A}{a-1}$	$\sigma^2 + \frac{bJ}{a-1} \sum_{i=1}^a \alpha_i^2$
Fator B	$b-1$	$\frac{1}{aJ} \sum_{k=1}^b B_k^2 - C$	$\frac{SQ_B}{b-1}$	$\sigma^2 + \frac{aJ}{b-1} \sum_{k=1}^b \beta_k^2$
Interação A×B	$(a-1)(b-1)$	$\frac{1}{J} \sum_{i=1}^a \sum_{k=1}^b (AB)_{ik} - C - SQ_A - SQ_B$	$\frac{SQ_{AB}}{(a-1)(b-1)}$	$\sigma^2 + \frac{J}{(a-1)(b-1)} \sum_{i=1}^a \sum_{k=1}^b (\alpha\beta)_{ik}^2$
Resíduo	$ab(J-1)$	Por diferença	$\frac{SQ_{RES}}{ab(J-1)}$	σ^2
Total	$abJ - 1$	$\sum_{i=1}^a \sum_{k=1}^b \sum_{j=1}^J x_{ijk}^2 - C$		

Teste F para o fator A :

$$F_A = \frac{\sigma^2 + \frac{bJ}{a-1} \sum_{i=1}^a \alpha_i^2}{\sigma^2}$$

Teste F para o fator B :

$$F_B = \frac{\sigma^2 + \frac{aJ}{b-1} \sum_{k=1}^b \beta_k^2}{\sigma^2}$$

Teste F para a existência de interação entre A e B :

$$F_{A \times B} = \frac{\sigma^2 + \frac{J}{(a-1)(b-1)} \sum_{i=1}^a \sum_{k=1}^b (\alpha\beta)_{ik}^2}{\sigma^2}$$

Exercício 6.2.1. Observar os dados de um ensaio inteiramente ao acaso, em fatorial 3×2 , no qual o objetivo é testar os efeitos de três recipientes para produção de mudas de duas espécies de *Eucalyptus*. A variável resposta medida foi a altura das mudas (em cm) aos 80 dias de idade. Os dados, vistos na tabela 6.2.2, foram adaptados do trabalho publicado por Simões (1970) e apresentados em Banzatto e Kronka (1989).

R1: Saco plástico pequeno

R2: Saco plástico grande

R3: Laminado

E1: *Eucalyptus citriodora*

E2: *Eucalyptus grandis*

Tabela 6.2.2. Altura das mudas de *Eucalyptus* (em cm) aos 80 dias de idade.

Tratamentos	Repetições				Total
	1	2	3	4	
R1E1	26,2	26,0	25,0	25,4	102,6
R1E2	24,8	24,6	26,7	25,2	101,3
R2E1	25,7	26,3	25,1	26,4	103,5
R2E2	19,6	21,1	19,0	18,6	78,3
R3E1	22,8	19,4	18,8	19,2	80,2
R3E2	19,8	21,4	22,8	21,3	85,3
Total					551,2

a) Análise de variância preliminar:

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - C = 26,2^2 + 26,0^2 + \dots + 21,3^2 - C = 198,7933$$

$$SQ_{TRAT} = \frac{1}{J} \sum_{i=1}^I T_i^2 - C = \frac{1}{4} (102,6^2 + 101,3^2 + \dots + 85,3^2) - C = 175,7033$$

$$C = \frac{G^2}{IJ} = \frac{551,2^2}{6 \times 4} = 12.659,2267$$

Tabela 6.2.3. ANOVA preliminar.

F.V.	G.L.	S.Q.	Q.M.
Tratamentos	5	175,7033	35,1407
Resíduo	18	23,0900	1,2828
Total	23	198,7933	

b) Obtenção das somas de quadrados dos fatores (R e E) e da interação ($R \times E$), efetuando o desdobramento dos cinco graus de liberdade de tratamentos.

Considerar o quadro auxiliar, apresentando os totais das seis combinações (tratamentos) e os totais dos níveis dos fatores *R* (Recipientes) e *E* (Espécies).

	R1	R2	R3	Total
E1	102,6	103,5	80,2	286,3
E2	101,3	78,3	85,3	264,9
Total	203,9	181,8	165,5	551,2

A figura 6.2.1 ilustra o comportamento das espécies E1 e E2 nos recipientes R1, R2 e R3.

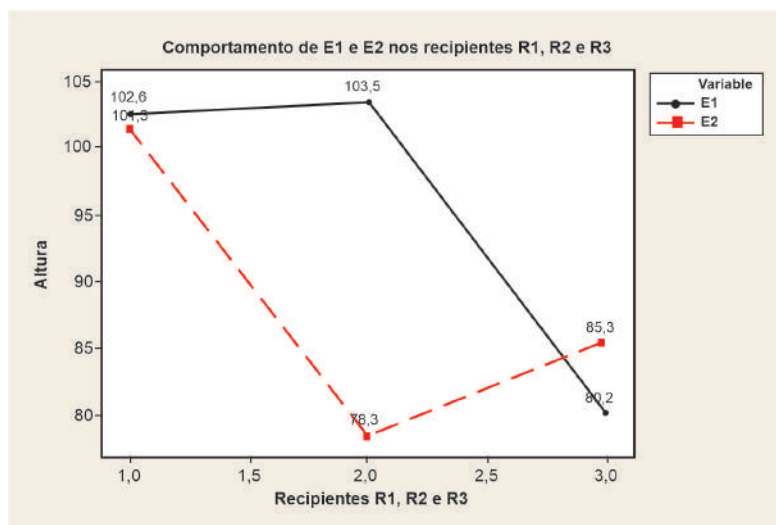


Figura 6.2.1. Comportamento de E1 e E2 nos recipientes R1, R2 e R3.

A figura 6.2.2 ilustra o comportamento dos recipientes R1, R2 e R3 em E1 e E2.

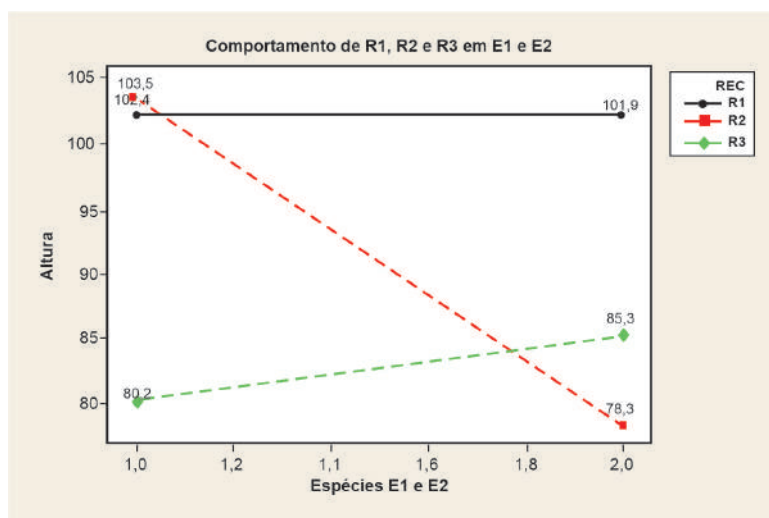


Figura 6.2.2. Efeito do comportamento de R1, R2 e R3 em E1 e E2.

Soma de quadrados de recipientes (SQ_R):

$$SQ_R = \frac{1}{8}(203,9^2 + 181,8^2 + 165,5^2) - C = 92,8608$$

Soma de quadrados de espécies (SQ_E):

$$SQ_E = \frac{1}{12}(286,3^2 + 264,9^2) - C = 19,0817$$

Soma de quadrados das combinações (tratamentos) ($SQ_{R,E}$):

$$SQ_{R,E} = \frac{1}{4}(102,6^2 + 103,5^2 + \dots + 85,3^2) - C = 175,7033$$

Soma de quadrados da interação ($SQ_{R \times E}$):

$$SQ_{R \times E} = SQ_{R,E} - SQ_R - SQ_E = 175,70 - 92,86 - 19,08 = 63,7608$$

Tabela 6.2.4. ANOVA com o desdobramento da $SQ_{TRAT.}$

F.V.	G.L.	S.Q.	Q.M.	F
Recipientes (R)	2	92,8608	46,4304	36,20**
Espécies (E)	1	19,0817	19,0817	14,88**
Interação (R×E)	2	63,7608	31,8804	24,85**
(Tratamentos)	(5)	(175,7033)		
Resíduo	18	23,0900	1,2828	
Total	23	198,7933		

$$F_{2;18;0,01} = 6,01 \quad F_{1;18;0,01} = 8,29$$

A interação entre os fatores recipiente e espécie é significativa, denotando existir dependência entre os fatores, demandando a necessidade de desdobrá-la, isto é, considerando cada fator dentro do outro.

Tabela 6.2.5. ANOVA no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Rec.	2	92.86083333	46.43041667	36.20	< 0.0001
Esp.	1	19,08166667	19,08166667	14,88	0,0012
Rec.×Esp.	2	63,76083333	31,88041667	24,85	< 0,0001
Resíduo	18	23,09000000	1,2827778		
Total	23	198.7933333			

Análise realizada pelo programa SAS, **Prog19**, em anexo no apêndice B.

c) Desdobramento da Interação $R \times E$ para estudar o comportamento das espécies dentro de cada recipiente:

	R1	R2	R3	Total
E1	102,6	103,5	80,2	286,3
E2	101,3	78,3	85,3	264,9
Total	203,9	181,8	165,5	551,2

$$SQ_{E d.R1} = \frac{1}{4}(102,6^2 + 101,3^2) - \frac{203,9^2}{8} = 0,2112$$

$$SQ_{E d.R2} = \frac{1}{4}(103,5^2 + 78,3^2) - \frac{181,8^2}{8} = 79,3800$$

$$SQ_{E d.R3} = \frac{1}{4}(80,2^2 + 85,3^2) - \frac{165,5^2}{8} = 3,2513$$

Observa-se que:

$$SQ_{E d.R1} + SQ_{E d.R2} + SQ_{E d.R3} = SQ_E + SQ_{R \times E} = 0,2112 + 79,3800 + 3,2513 = 19,0817 + 63,7608 = 82,8425$$

Tabela 6.2.6. ANOVA de espécies dentro de cada recipiente.

F.V.	G.L.	S.Q.	Q.M.	F
Recipientes (R)	2	92,8608	46,4304	36,20 **
Espécies d. R1	1	0,2112	0,2112	0,16 <i>ns</i>
Espécies d. R2	1	79,3800	79,3800	61,88 **
Espécies d. R3	1	3,2513	3,2513	2,53 <i>ns</i>
(Tratamentos)	(5)	(175,7033)		
Resíduo	18	23,0900	1,2828	
Total	23	198,7933		

$$F_{1;18;0,05} = 4,41 \quad F_{1;18;0,01} = 8,29$$

Desdobramento de espécies dentro de cada recipiente					
Rec.	G.L.	S.Q.	Q.M.	F	Pr > F
R1	1	0,211250	0,211250	0,16	0,6897
R2	1	79,380000	79,380000	61,88	< 0,0001
R3	1	3,251250	3,251250	2,53	0,1288

A análise permite concluir que o saco plástico pequeno (R1) ou o laminado (R3) não produziram diferença significativa no desenvolvimento das mudas das duas espécies (p -valor > 0,05).

Por outro lado, pode-se concluir que o saco plástico grande (R2) produziu diferença significativa no desenvolvimento das mudas das duas espécies (p -valor < 0,01).

Os resultados dos testes F confirmam as tendências mostradas na figura 6.2.1, que ilustra o comportamento de E1 e E2 nos recipientes R1, R2 e R3.

d) Desdobramento da Interação $R \times E$ para estudar o comportamento dos recipientes dentro de cada espécie:

$$SQ_{R.d.E1} = \frac{1}{4}(102,6^2 + 103,5^2 + 80,2^2) - \frac{286,3^2}{12} = 87,1217$$

$$SQ_{R.d.E2} = \frac{1}{4}(101,3^2 + 78,3^2 + 85,3^2) - \frac{264,9^2}{12} = 69,5000$$

Em que:

$$SQ_{R.d.E1} + SQ_{R.d.E2} = SQ_R + SQ_{R \times E} = 87,1217 + 69,5000 = 92,86083 + 63,76083 = 156,6217$$

Tabela 6.2.7. ANOVA de recipientes dentro de cada espécie.

F.V.	G.L.	S.Q.	Q.M.	F
Espécies (E)	1	19,0817	19,0817	14,88 **
Recipientes d. E1	2	87,1217	43,5608	33,96 **
Recipientes d. E2	2	69,5000	34,7500	27,09 **
(Tratamentos)	(5)	(175,7033)		
Resíduo	18	23,0900	1,2828	
Total	23	198,7933		

$$F_{2;18;0,01} = 6,01$$

Há razão para concluir que os três recipientes possuem efeitos diferentes sobre o desenvolvimento das mudas de *Eucalyptus* (p -valor $< 0,01$).

Desdobramento de recipientes dentro de cada espécie					
Esp.	G.L.	S.Q.	Q.M.	F	Pr > F
E1	2	87,121667	43,560833	33,96	< 0,0001
E2	2	69,500000	34,750000	27,09	< 0,0001

Os resultados dos testes F confirmam as tendências mostradas na figura 6.2.2, a qual ilustra o comportamento de R1, R2 e R3 dentro das espécies E1 e E2.

e) Teste de comparação múltipla de Tukey

e1) Recipientes dentro de E1 (*Eucalyptus citriodora*)

$$\Delta = q_{3;18;0,05} \sqrt{\frac{QM_{RES}}{J}} = 3,61 \sqrt{\frac{1,2828}{4}} = 2,0444 \quad q_{3;18;0,05} = 3,61$$

$$\overline{R2E1} = 103,5 / 4 = 25,88...a$$

$$\overline{R1E1} = 102,6 / 4 = 25,65...a$$

$$\overline{R3E1} = 80,2 / 4 = 20,05....b$$

Conclui-se que o desenvolvimento das mudas de *Eucalyptus citriodora* (E1) apresentou comportamento similar nos recipientes R1 e R2, pois não houve diferença significativa entre eles. Entretanto, os recipientes R1 e R2 apresentaram diferenças significativas em relação ao recipiente R3 (laminado), não sendo recomendada a utilização do laminado para a produção das mudas da espécie E1 (*Eucalyptus citriodora*).

e2) Recipientes dentro de E2 (*Eucalyptus grandis*)

$$\Delta = q_{3;18;0,05} \sqrt{\frac{QM_{RES}}{J}} = 3,61 \sqrt{\frac{1,2828}{4}} = 2,0444 \quad q_{3;18;0,05} = 3,61$$

$$\overline{R1E2} = 25,33...a$$

$$\overline{R2E2} = 19,58...b$$

$$\overline{R3E2} = 21,33...b$$

Há razão para concluir, no desenvolvimento das mudas, que o *Eucalyptus grandis* (E2) apresentou comportamento similar nos recipientes R2 (saco grande) e R3 (laminado), pois não houve diferença significativa entre os mesmos. Entretanto, o recipiente R1 (saco pequeno) apresentou diferença significativa em relação aos recipientes R2 (saco grande) e R3 (laminado), concluindo-se ser recomendada a utilização do saco pequeno na produção das mudas da espécie E2 (*Eucalyptus grandis*).

6.3. EXPERIMENTOS FATORIAIS HIERÁRQUICOS

Nos delineamentos experimentais, as combinações que envolvem dois ou mais fatores podem gerar dois tipos distintos de classificação: cruzada e hierárquica (aninhada).

Na classificação cruzada, cada nível de um fator combina com cada nível dos demais e o resultado é um cruzamento completo entre os fatores. No exemplo de um experimento, considerando quatro espécies florestais (E) e três tipos de tratamentos silviculturais (T), são geradas 12 combinações, ilustradas no quadro a seguir:

E1T1	E2T1	E3T1	E4T1
E1T2	E2T2	E3T2	E4T2
E1T3	E2T3	E3T3	E4T3

É possível, estabelecer os seguintes efeitos dos fatores na classificação cruzada:

a) Efeito principal: é o efeito produzido em cada fator, mas não considerando os níveis dos demais fatores;

b) Efeito da interação: reflete o comportamento de cada fator, considerando os níveis dos demais fatores, representado conforme o exemplo apresentado por E×T. A análise da interação E×T avaliará se o comportamento das quatro espécies corresponde ao mesmo, levando em consideração os três tratamentos silviculturais e vice-versa. Por outro lado, se comprovada a ocorrência de um mesmo comportamento em todos os níveis, então a conclusão será de não existência de interação entre os fatores E e T.

A ocorrência de interação significativa não permite interpretar os efeitos principais dos fatores. Nesse caso, a interação deve ser desdobrada e a análise de cada fator realizada, separadamente, estudando o comportamento de cada fator em cada nível do outro. Assim sendo, devem-se estudar os desdobramentos E d. T1, E d. T2, E d. T3, o que corresponde à verificação do comportamento das espécies (E), independentemente, em cada tratamento silvicultural (T).

Em um experimento aninhado, comparam-se níveis de fatores principais e níveis de fatores secundários, que variam “dentro” de cada um dos fatores principais. Os tratamentos secundários não se repetem para os diversos tratamentos principais. Dentro de cada tratamento principal, há uma ramificação em tratamentos secundários que são específicos, isto é, os tratamentos secundários.

Em um experimento inteiramente ao acaso com quatro repetições, deseja-se testar quatro espécies do gênero *Eucalyptus* (fator *A*) versus quatro procedências dentro de cada espécie (Fator *B*), correspondendo ao croqui de um experimento hierárquico em dois estágios.

E1	Procedência 11	E2	Procedência 21
	Procedência 12		Procedência 22
	Procedência 13		Procedência 23
	Procedência 14		Procedência 24
E3	Procedência 31	E4	Procedência 41
	Procedência 32		Procedência 42
	Procedência 33		Procedência 43
	Procedência 34		Procedência 44

Neste experimento, os níveis do fator *B* (procedência) apresentam similaridades para os diferentes níveis, mas não são idênticos. Essa estrutura de arranjo dos níveis dos fatores (tratamentos) origina os denominados delineamentos hierárquicos ou de aninhamento, haja vista que os níveis do fator *B* são hierárquicos (estão aninhados) em relação aos níveis do fator *A* (espécies). No exemplo citado, a combinação E2P23, segundo nível de *A*, representada por E2, com o terceiro nível de *B*, simbolizada por P23, indica que P23 não ocorrerá com qualquer outro nível do fator *A* (*E*), o que caracteriza uma relação de aninhamento.

Essa relação é denotada por $B(A)$, em que *B* é dito fator aninhado e *A* chamado fator de aninhamento. Diferentemente da classificação cruzada, a relação hierárquica não é recíproca, ou seja, $B(A)$ é diferente de $A(B)$. Assim sendo, nos delineamentos de classificação hierárquica, não é possível avaliar a interação entre os fatores *A* e *B*, mas é estimado o efeito médio dos níveis do fator *B* dentro dos níveis do fator *A*, isto é, no caso dos delineamentos hierárquicos com dois fatores que apresentam dois estágios. No caso de três fatores, *A*, *B* e *C*, pode-se ter o estudo dos efeitos médios dos níveis de *C* dentro de *B* e *A*; e os de *B* dentro de *A*, assim sucessivamente.

6.3.1. Análise de variância de experimentos hierárquicos em dois estágios

Considerar a análise de variância para um fatorial hierárquico balanceado em dois estágios (fatores *A* e *B*), em que o balanceamento é garantido por igual número de níveis de *B* dentro de cada nível

de A , bem como pelo número de repetições constante para todos os tratamentos (combinações de níveis).

Observar o modelo matemático, admitindo o delineamento experimental inteiramente ao acaso, considerando o exemplo citado sobre espécies (E) e procedências (P) aninhadas em espécie.

$$x_{ikj} = m + sp_i + p_{ik} + e_{ikj} \text{ ou } x_{ikj} = m + sp_i + p_{k(i)} + e_{(ik)j}$$

$k(i)$: significa que o nível k do fator B está aninhado sob o nível i do fator A ;

$i = 1, \dots, a$: níveis do fator A (Espécies);

$k = 1, \dots, b$; níveis do fator B (Procedências) dentro de cada nível de A ;

$j = 1, \dots, J$: número de repetições.

É recomendável se referir às repetições como dentro de uma combinação de níveis de A e B . Assim sendo, o índice $(ik)j$ é usado para o termo do erro. As tabelas 6.3.1.1 e 6.3.1.2 apresentam, respectivamente, o quadro da análise de variância e os componentes de variância para o ensaio inteiramente ao acaso, considerando os efeitos fixos e aleatórios para as espécies e procedências.

Tabela 6.3.1.1. ANOVA do fatorial hierárquico em dois estágios.

F.V.	G.L.	S.Q.	Q.M.	F
Entre espécies	$a - 1$	SQ_A	$SQ_A / (a - 1)$	$QM_A / QM_{B(A)}$
Entre procedências d. espécie	$a(b - 1)$	$SQ_{B(A)}$	$SQ_{B(A)} / a(b - 1)$	$QM_{B(A)} / QM_{RES}$
Resíduo	$ab(J - 1)$	$SQ_{RESIDUO}$	$SQ_{RES} / ab(J - 1)$	
Total	$abJ - 1$	SQ_{TOTAL}		

sp_i : efeito de espécie considerado aleatório ($i = 1, 2, \dots, a$), então $NID(0, \sigma_{sp}^2)$, em que $E(sp_i) = 0$ e $E(sp_i^2) = \sigma_{sp}^2$;

sp_i : efeito de espécie considerado fixo ($i = 1, 2, \dots, a$), em que $E(sp_i) = sp_i$, $E(sp_i^2) = sp_i^2$ e $\sum_{i=1}^a sp_i = 0$;

p_{ik} : efeito de procedência k dentro de espécie i considerado aleatório ($k = 1, 2, \dots, b$), em que $NID(0, \sigma_p^2)$, $E(p_{ik}) = 0$ e $E(p_{ik}^2) = \sigma_p^2$;

p_{ik} : efeito de procedência k dentro de espécie i considerado fixo ($k = 1, 2, \dots, b$), em que $E(p_{ik}) = p_{ik}$, $E(p_{ik}^2) = p_{ik}^2$ e $\sum_{i=1}^a \sum_{k=1}^b p_{i(k)} = 0$;

e_{ikj} : é o erro experimental, tal que, $NID(0, \sigma^2)$, ou seja, $E(e_{ijk}) = 0$ e $E(e_{ijk}^2) = \sigma^2$.

Tabela 6.3.1.2. Componentes de variância do fatorial hierárquico em dois estágios.

$E(QM)$	A e B fixos	A fixo e B aleatório	A e B aleatórios
$E(QM_A)$	$\sigma^2 + \frac{bJ \sum_{i=1}^a sp_i^2}{a-1}$	$\sigma^2 + J\sigma_p^2 + \frac{bJ \sum_{i=1}^a sp_i^2}{a-1}$	$\sigma^2 + J\sigma_p^2 + bJ\sigma_{sp}^2$
$E(QM_{B(A)})$	$\sigma^2 + \frac{J \sum_{i=1}^a \sum_{k=1}^b p_{k(i)}^2}{a(b-1)}$	$\sigma^2 + J\sigma_p^2$	$\sigma^2 + J\sigma_p^2$
$E(QM_{RES})$	σ^2	σ^2	σ^2

Considerando as abJ unidades experimentais, ocorre a decomposição da variação total de um fatorial hierárquico em dois estágios:

$$SQ_{TOTAL} = \sum_{i=1}^a \sum_{k=1}^b \sum_{j=1}^J x_{ikj}^2 - C \quad SQ_A = \frac{1}{bJ} \sum_{i=1}^a x_{i..}^2 - C \quad C = \frac{(\sum_{i=1}^a \sum_{k=1}^b \sum_{j=1}^J x_{ikj})^2}{abJ}$$

$$SQ_{B(A)} = \sum_{i=1}^a \sum_{k=1}^b x_{k(i).}^2 - \frac{1}{bJ} \sum_{i=1}^a x_{i..}^2 \quad SQ_{RES} = \sum_{i=1}^a \sum_{k=1}^b \sum_{j=1}^J x_{ikj}^2 - \frac{1}{J} \sum_{i=1}^a \sum_{k=1}^b x_{k(i).}^2$$

Os graus de liberdade $a(b-1)$ para $B(A)$ representam a soma dos graus de liberdade de B e da interação $A \times B$, ou seja, $a(b-1) = (b-1) + (a-1)(b-1)$. Por outro lado, caso se tratasse de um fatorial cruzado, haveria, separadamente, os graus de liberdade do fator B e da interação $A \times B$.

A ocorrência de significância estatística do fator A indica ocorrer diferenciação nas respostas médias produzidas pelos diferentes níveis do fator A . No caso da fonte de variação $B(A)$, a significância estatística indica que as respostas médias entre níveis de B , para um mesmo nível de A , mostram efeitos diferenciados. Essas investigações podem explicar se as causas da variabilidade das observações estão preponderantemente entre os níveis de A , ou dentro desses níveis, como resultado do efeito do fator B .

Os quadrados médios (Q.M.) são obtidos dividindo-se cada soma de quadrados pelos seus respectivos graus de liberdade. Considerando que os erros são normais e independentemente distribuídos $(0, \sigma^2)$, a razão entre quaisquer dois quadrados médios possui distribuição F de Snedecor.

Na tabela 6.3.1.3, de acordo com Albuquerque (2016), estão os componentes de variância para um experimento aninhado em três estágios, considerando A e B fixos e C aleatório.

$$x_{iklj} = m + \tau_i + \beta_{k(i)} + \gamma_{l(ik)} + e_{j(ikl)}$$

$$i = 1, 2, \dots, a \quad k = 1, 2, \dots, b \quad l = 1, 2, \dots, c \quad j = 1, 2, \dots, J$$

Tabela 6.3.1.3. Componentes de variância do fatorial hierárquico em três estágios.

$E(QM)$	$A \text{ e } B \text{ fixos e } C \text{ aleatório}$
$E(QM_A)$	$\sigma^2 + j\sigma_\gamma^2 + \frac{bcJ \sum_{i=1}^a \tau_i^2}{a-1}$
$E(QM_{B(A)})$	$\sigma^2 + J\sigma_\gamma^2 + \frac{cJ \sum_{i=1}^a \sum_{k=1}^b \beta_{k(i)}^2}{a(b-1)}$
$E(QM_{C(B(A))})$	$\sigma^2 + J\sigma_\gamma^2$
$E(QM_{RES})$	σ^2

Exercício 6.3.1.1. Considerando um experimento inteiramente ao acaso com quatro repetições, deseja-se testar quatro espécies do gênero *Eucalyptus* (E : fator A) versus quatro procedências (P) dentro de cada espécie (P : Fator B), correspondendo ao croqui de um experimento hierárquico em dois estágios. Os níveis do fator A e do fator B são fixos. A tabela 6.3.1.3 apresenta os dados para a análise.

Tabela 6.3.1.3. Produção de madeira (dados fictícios) em $m^3 \text{ ha}^{-1}$.

E	P	Prod.	E	P	Prod.	E	P	Prod.	E	P	Prod.
E1	P1	177,1	E1	P2	166,2	E1	P3	190,6	E1	P4	169,1
E1	P1	187,3	E1	P2	174,7	E1	P3	170,9	E1	P4	197,9
E1	P1	186,8	E1	P2	183,9	E1	P3	174,0	E1	P4	205,1
E1	P1	172,8	E1	P2	168,5	E1	P3	189,5	E1	P4	186,1
E2	P5	202,8	E2	P6	196,8	E2	P7	209,6	E2	P8	195,6
E2	P5	200,2	E2	P6	205,3	E2	P7	209,1	E2	P8	207,0
E2	P5	207,9	E2	P6	206,6	E2	P7	203,2	E2	P8	203,6
E2	P5	190,1	E2	P6	194,7	E2	P7	197,3	E2	P8	204,9
E3	P9	187,2	E3	P10	203,7	E3	P11	201,2	E3	P12	190,6
E3	P9	197,8	E3	P10	193,6	E3	P11	196,5	E3	P12	196,1
E3	P9	191,9	E3	P10	194,4	E3	P11	197,8	E3	P12	202,9
E3	P9	199,5	E3	P10	197,8	E3	P11	191,5	E3	P12	202,5
E4	P13	193,6	E4	P14	206,3	E4	P15	186,0	E4	P16	172,2
E4	P13	204,2	E4	P14	202,0	E4	P15	172,2	E4	P16	194,0
E4	P13	206,3	E4	P14	176,1	E4	P15	187,5	E4	P16	187,8
E4	P13	202,0	E4	P14	173,5	E4	P15	179,0	E4	P16	200,1

a) Soma de quadrados para o total:

$$SQ_{TOTAL} = \sum_{i=1}^a \sum_{k=1}^b \sum_{j=1}^J x_{ikj}^2 - C$$

$$C = \frac{(\sum_{i=1}^a \sum_{k=1}^b \sum_{j=1}^J x_{ikj})^2}{abJ} = \frac{12.323,0^2}{4 \times 4 \times 4} = 2.372.755,1406$$

$$SQ_{TOTAL} = 2.381.787,36 - C = 9.032,2194$$

b) Soma de quadrados para espécies (A):

Considerar o quadro auxiliar abaixo:

E1		E2		E3		E4	
P1	724,0	P5	801,0	P9	776,4	P13	806,1
P2	693,3	P6	803,4	P10	789,5	P14	757,9
P3	725,0	P7	819,2	P11	787,0	P15	724,7
P4	758,2	P8	811,1	P12	792,1	P16	754,1
Total	2.900,5	Total	3.234,7	Total	3.145,0	Total	3.042,8

$$SQ_A = \frac{1}{bJ} \sum_{i=1}^a x_{i..}^2 - C$$

$$SQ_A = \frac{1}{4 \times 4} (2.900,5^2 + 3.234,7^2 + 3.145^2 + 3.042,8^2) - C = 3.859,9331$$

$$SQ_A = \frac{1}{4 \times 4} (9.512.322,68) - C = 3.859,9331$$

c) Soma de quadrados para procedências (B) aninhadas em espécies (A):

$$SQ_{A(B)} = \frac{1}{J} \sum_{i=1}^a \sum_{k=1}^b x_{k(i).}^2 - \frac{1}{bJ} \sum_{i=1}^a x_{i..}^2$$

$$SQ_{A(B)} = \frac{1}{4} (724,0^2 + 693,3^2 + \dots + 724,7^2 + 754,1^2) - \frac{1}{4 \times 4} (2.900,5^2 + \dots + 3.042,8^2) = 1.465,5962$$

d) Soma de quadrados para o resíduo:

$$SQ_{RES} = \sum_{i=1}^a \sum_{k=1}^b \sum_{j=1}^J x_{ikj}^2 - \frac{1}{J} \sum_{i=1}^a \sum_{k=1}^b x_{k(i).}^2$$

$$SQ_{RES} = 2.381.787,36 - \frac{1}{4} 9.512.322,68 = 3.706,69$$

e) Quadro da análise de variância:

Tabela 6.3.1.4. ANOVA.

F.V	G.L.	S.Q	M.Q.	F
Espécies (A)	3	3.859,9331	1.286,6443	10,53 **
Procedências (B) d. Espécies	12	1.465,5962	122,1330	1,58 ns
Resíduo	48	3.706,6900	77,2227	
Total	63	9.032,2193		

Tabela 6.3.1.5. ANOVA no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Espécies	3	3859,933125	1286,644375	16,66	< 0,0001
Procedências (Espécies)	12	1465,596250	122,133021	1,58	0,1292
Resíduo	48	3706,690000	77,222708		
Total	63	9032,219375			

Teste de hipótese usando Procedência (Espécie) como resíduo					
F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Espécies	3	3859,933125	1286,644375	10,53	0,0011

O programa **Prog20**, anexo no apêndice B, efetua a análise do delineamento hierárquico inteiramente ao acaso em dois estágios. O comando “Test h = Espécie e = Procedencia (Espécie);” corrige os valores do teste *F* referente ao fator *A* (Espécie), em que o seu quadrado médio deve ser dividido pelo quadrado médio de procedência d. Espécie.

f) Teste de comparação múltipla de Tukey para o fator *A* (espécie):

$$\Delta = q_{4; 12; 0,05} \times \sqrt{\frac{QM_{P d E}}{16}} = 4,19852 \sqrt{\frac{122,1330}{16}} = 11,6$$

Espécies	Médias	
E2	202,2	a
E3	196,6	ab
E4	190,2	bc
E1	181,3	c

As médias seguidas pela mesma letra não diferem pelo teste de Tukey, ao nível de probabilidade $\alpha = 0,05$. As espécies E2 e E3 apresentaram os melhores desempenhos.

Alfa (α)	0.05
Graus de liberdade do resíduo	12
Quadrado médio do resíduo	122,133
Valor crítico da amplitude total estudentizada	4,19852
Diferença mínima de significância	11,6

Tukey (Agrupamento)		Médias	N	Espécie
	A	202,169	16	E2
B	A	196,563	16	E3
B	C	190,175	16	E4
	C	181,281	16	E1

g) Desdobramento das procedências (B) dentro das espécies (A):

E1		E2		E3		E4	
P1	724,0	P5	801,0	P9	776,4	P13	806,1
P2	693,3	P6	803,4	P10	789,5	P14	757,9
P3	725,0	P7	819,2	P11	787,0	P15	724,7
P4	758,2	P8	811,1	P12	792,1	P16	754,1
Total	2.900,5	Total	3.234,7	Total	3.145,0	Total	3.042,8

$$SQ_{P.d.E1} = \frac{1}{4}(724,0^2 + 693,3^2 + 725,0^2 + 758,2^2) - \frac{2.900,5^2}{16} = 527,0169$$

$$SQ_{P.d.E2} = \frac{1}{4}(801,0^2 + 803,4^2 + 819,2^2 + 811,1^2) - \frac{3.234,7^2}{16} = 50,8469$$

$$SQ_{P.d.E3} = \frac{1}{4}(776,4^2 + 789,5^2 + 787,0^2 + 792,1^2) - \frac{3.145,0^2}{16} = 35,5925$$

$$SQ_{P.d.E4} = \frac{1}{4}(806,1^2 + 757,9^2 + 724,7^2 + 754,1^2) - \frac{3.042,8^2}{16} = 852,1400$$

Tabela 6.3.1.6. ANOVA de procedências dentro de espécies no SAS.

Desdobramento das procedências dentro das espécies					
Espécies	G.L.	S.Q.	Q.M.	F	Pr > F
E1	3	527,016875	175,672292	2,27	0,0918
E2	3	50,846875	16,948958	0,22	0,8824
E3	3	35,592500	11,864167	0,15	0,9268
E4	3	852,140000	284,046667	3,68	0,0183
Resíduo	48	3706,690000	77,222708		

$$SQ_{P.d.E} = SQ_{P.d.E1} + SQ_{P.d.E2} + SQ_{P.d.E3} + SQ_{P.d.E4} = 1.465,5963$$

h) Teste de comparação múltipla de Tukey para as procedências (B) dentro da espécie E4:

Considerando que, somente as procedências dentro da espécie E4 apresentaram significância ao nível de significância $\alpha = 0,05$, o teste de Tukey será aplicado somente para as procedências P13, P14, P15 e P16.

$$\Delta = q_{4; 48; 0,05} \times \sqrt{\frac{QM_{RES}}{4}} = 3,77 \sqrt{\frac{77,2227}{4}} = 16,5647$$

Espécie4		
Procedências	Médias	
P13	201,5250	a
P14	189,4750	ab
P16	188,5250	b
P15	181,1750	b

É recomendável utilizar a procedência P13 ou P14, pois apresentaram o melhor comportamento dentro da espécie E4.

6.4. CONFUNDIMENTO

Nos experimentos fatoriais com elevado número de tratamentos, especialmente em ensaios de adubação, em que o tamanho da parcela é relativamente grande podendo comprometer a homogeneidade dos blocos, a técnica de confundimento visa diminuir o número de tratamentos por bloco. Isso é feito dividindo, segundo determinada orientação, um bloco em dois ou mais sub-blocos, objetivando garantir a homogeneidade do material experimental, sem comprometer o balanceamento que existe nesses experimentos.

A técnica do confundimento resume-se em fazer com que um dos efeitos, ou parte dele, seja igual ao efeito de blocos. Como geralmente acontece em ensaios com três fatores em que a interação tripla é não significativa, então é realizado o confundimento da interação tripla com blocos. Na análise estatística, dois efeitos são denominados confundidos quando não é possível isolar o efeito de cada um deles.

6.4.1. Confundimento do fatorial 2^3

Considerar um experimento de adubação de uma espécie florestal em fatorial 2^3 , no qual as condições do terreno não permitem estabelecer blocos com oito parcelas. Dessa maneira, é possível planejar o confundimento da interação tripla com blocos, resultando em oito blocos incompletos com quatro parcelas cada um. Nesse caso, existe um confundimento dos blocos com a interação $N \times P \times K$, ou seja, o contraste estima ao mesmo tempo a interação $N \times P \times K$ e a diferença entre os blocos. Então, não se pode estimar bem nenhum desses efeitos.

A seguir o quadro dos contrastes do fatorial 2^3 com $N \times P \times K$:

	$N_0P_0K_0$	$N_1P_0K_0$	$N_0P_1K_0$	$N_0P_0K_1$	$N_1P_1K_0$	$N_1P_0K_1$	$N_0P_1K_1$	$N_1P_1K_1$
N	-	+	-	-	+	+	-	+
P	-	-	+	-	+	-	+	+
K	-	-	-	+	-	+	+	+
$N \times P$	+	-	-	+	+	-	-	+
$N \times K$	+	-	+	-	-	+	-	+
$P \times K$	+	+	-	-	-	-	+	+
$N \times P \times K$	-	+	+	+	-	-	-	+

Considerar o contraste da interação tripla para a constituição dos blocos incompletos de acordo com o quadro auxiliar:

$$Y_{N \times P \times K} = -N_0P_0K_0 + N_1P_0K_0 + N_0P_1K_0 + N_0P_0K_1 - N_1P_1K_0 - N_1P_0K_1 - N_0P_1K_1 + N_1P_1K_1$$

Observando que a interação $N \times P \times K$, em geral, é não significativa, é necessário adotar, como sub-bloco1, os tratamentos que aparecem na interação tripla com os sinais positivos; e, como sub-bloco2, os que aparecem com sinais negativos.

Sub-bloco1(+)	Sub-bloco2 (-)
$N_1P_0K_0$	$N_0P_0K_0$
$N_0P_1K_0$	$N_1P_1K_0$
$N_0P_0K_1$	$N_1P_0K_1$
$N_1P_1K_1$	$N_0P_1K_1$

Os blocos originais serão chamados de repetições e os sub-blocos de blocos, como se fossem blocos incompletos. Logo, em cada repetição, haverá dois sub-blocos: um sub-bloco com os tratamentos $N_1P_0K_0$, $N_0P_1K_0$, $N_0P_0K_1$ e $N_1P_1K_1$ distribuídos ao acaso dentro do sub-bloco1; e o outro, por sua vez, com os tratamentos $N_0P_0K_0$, $N_1P_1K_0$, $N_1P_0K_1$ e $N_0P_1K_1$ também distribuídos ao acaso dentro do sub-bloco2.

É importante observar que cada fator, em cada sub-bloco (bloco incompleto), ocorre duas vezes em cada um dos níveis. Assim sendo, a metodologia usada mantém o balanceamento do experimento.

Por exemplo, no caso de 4 repetições, há 4 sub-blocos1 e 4 sub-blocos2, totalizando 8 sub-blocos (blocos), então:

$$SQ_{BLOCOS} = \frac{1}{4}(B_1^2 + \dots + B_8^2) - C$$

Observar a análise de variância de um fatorial 2^3 com quatro repetições (em blocos) com confundimento. Como existem 4 repetições (blocos), sendo cada repetição com 2 sub-blocos, a análise com confundimento terá 8 sub-blocos.

Para facilitar o entendimento, a tabela 6.4.1.1 apresenta a ANOVA do fatorial 2^3 com e sem confundimento.

Tabela 6.4.1.1. ANOVA do fatorial 2^3 com e sem confundimento.

F.V.	G.L. (Sem confundimento.)	G.L. (Com confundimento.)
N	1	1
P	1	1
K	1	1
$N \times P$	1	1
$N \times K$	1	1
$P \times K$	1	1
$N \times P \times K$	1	-----
(Tratamentos)	(7)	(6)
Blocos	3	7
Resíduo	21	18
Total	31	31

Na análise com confundimento, a interação $N \times P \times K$ deixou de existir e o número de blocos (blocos incompletos) passou para oito.

Por outro lado, caso houvesse o interesse em utilizar uma interação dupla, em vez da interação tripla, como exemplo a interação $N \times P$, a técnica seria a mesma.

Para uma melhor compreensão, um ensaio com adubação de um fatorial 2^3 é apresentado abaixo. Ele considera a produção de madeira de uma espécie florestal em m^3 por parcela, com quatro blocos, considerando os fatores N , P , K (ausência e presença).

	$N_0P_0K_0$	$N_1P_0K_0$	$N_0P_1K_0$	$N_1P_1K_0$	$N_0P_0K_1$	$N_1P_0K_1$	$N_0P_1K_1$	$N_1P_1K_1$
B1	17,6	23,0	21,4	22,1	31,6	33,2	28,1	35,8
B2	26,5	27,4	32,5	45,8	42,5	38,5	26,1	39,4
B3	22,4	35,8	22,2	32,1	34,8	28,0	24,6	29,6
B4	29,1	35,9	31,6	36,3	33,5	33,5	33,0	40,2

A tabela 6.4.1.2 mostra os resultados da análise de variância do ensaio com adubação de um fatorial 2^3 , sem o uso da técnica de confundimento.

Tabela 6.4.1.2. ANOVA fatorial 2^3 sem uso do confundimento.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	394,859375	131,3619792	6,35	0,0031
N	1	195,5253125	195,5253125	9,46	0,0057
P	1	1,7578125	1,7578125	0,09	0,7735
K	1	156,2028125	156,2028125	7,55	0,0120
$N \times P$	1	61,8828125	61,8828125	2,99	0,0983

Tabela 6.4.1.2. ANOVA fatorial 2^3 sem uso do confundimento (continuação).

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
N×K	1	30,2253125	30,2253125	1,46	0,2401
P×K	1	63,5628125	63,5628125	3,07	0,0941
N×P×K	1	50,7528125	50,7528125	2,45	0,1321
Resíduo	21	434,236563	20,677932		
Total	31	1388,232188			

Por outro lado, adotando, como sub-bloco1, os tratamentos que aparecem na interação tripla com os sinais positivos e, como sub-bloco2, os que aparecem com o sinal negativo, ocorrem, na tabela 6.4.1.3, os resultados da análise de variância com o confundimento da interação tripla com blocos.

B1	$N_0P_0K_0$	$N_1P_1K_0$	$N_1P_0K_1$	$N_0P_1K_1$	B5	$N_0P_0K_0$	$N_1P_1K_0$	$N_1P_0K_1$	$N_0P_1K_1$
B2	$N_1P_0K_0$	$N_0P_1K_0$	$N_0P_0K_1$	$N_1P_1K_1$	B6	$N_1P_0K_0$	$N_0P_1K_0$	$N_0P_0K_1$	$N_1P_1K_1$
B3	$N_0P_0K_0$	$N_1P_1K_0$	$N_1P_0K_1$	$N_0P_1K_1$	B7	$N_0P_0K_0$	$N_1P_1K_0$	$N_1P_0K_1$	$N_0P_1K_1$
B4	$N_1P_0K_0$	$N_0P_1K_0$	$N_0P_0K_1$	$N_1P_1K_1$	B8	$N_1P_0K_0$	$N_0P_1K_0$	$N_0P_0K_1$	$N_1P_1K_1$

Distribuídos os tratamentos ao acaso dentro de cada sub-bloco, tem-se os valores para a variável resposta para os oito sub-blocos:

B1	17,6	22,1	33,2	28,1	B5	22,4	32,1	28,0	24,6
B2	23,0	21,4	31,6	35,8	B6	35,8	22,2	34,8	29,6
B3	26,5	45,8	38,5	26,1	B7	29,1	36,3	33,5	33,0
B4	27,4	32,5	42,5	39,4	B8	35,9	31,6	33,5	40,2

Observa-se, na análise da tabela 6.4.1.3, a ausência da interação $N \times P \times K$, a qual foi confundida com os blocos.

O uso do confundimento produz sempre uma perda no número de graus de liberdade do resíduo.

Comparando as tabelas 6.4.1.2 e 6.4.1.3, verifica-se que houve uma redução de três graus de liberdade, quando realizado o confundimento, no número de graus de liberdade do resíduo.

Os sete graus de liberdade de blocos, na análise com confundimento, correspondem à soma dos três graus de liberdade da análise sem confundimento, um da interação $N \times P \times K$ e os três graus de liberdade que foram retirados do resíduo.

Observa-se (tabelas 6.4.1.2 e 6.4.1.3) que os efeitos do nitrogênio (N) e do potássio (K) foram significativos, enquanto os efeitos do fósforo (P) e das interações $N \times P$, $N \times K$ e $P \times K$ foram não significativos, ao nível $\alpha = 0,05$ de significância.

Tabela 6.4.1.3. ANOVA fatorial 2^3 com o uso do confundimento.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	7	451,7396875	64,5342411	2,72	0,0413
N	1	195,5253125	195,5253125	8,24	0,0102
P	1	1,7578125	1,7578125	0,07	0,7886
K	1	156,2028125	156,2028125	6,58	0,0195
N×P	1	61,8828125	61,8828125	2,61	0,1238
N×K	1	30,2253125	30,2253125	1,27	0,2740
P×K	1	63,5628125	63,5628125	2,68	0,1191
Resíduo	18	427,335625	23,740868		
Total	31	1388,232187			

6.4.2. Confundimento do fatorial 3^3

Os experimentos fatoriais $3 \times 3 \times 3$ ou 3^3 consistem na aplicação de três fatores, cada um com três níveis, resultando em um conjunto de 27 tratamentos. No caso de um experimento com adubação NPK, apresentando cada fator níveis equidistantes e representados por 0, 1, 2, tem-se:

$N_0P_0K_0$	000	$N_1P_0K_0$	100	$N_2P_0K_0$	200
$N_0P_0K_1$	001	$N_1P_0K_1$	101	$N_2P_0K_1$	201
$N_0P_0K_2$	002	$N_1P_0K_2$	102	$N_2P_0K_2$	202
$N_0P_1K_0$	010	$N_1P_1K_0$	110	$N_2P_1K_0$	210
$N_0P_1K_1$	011	$N_1P_1K_1$	111	$N_2P_1K_1$	211
$N_0P_1K_2$	012	$N_1P_1K_2$	112	$N_2P_1K_2$	212
$N_0P_2K_0$	020	$N_1P_2K_0$	120	$N_2P_2K_0$	220
$N_0P_2K_1$	021	$N_1P_2K_1$	021	$N_2P_2K_1$	221
$N_0P_2K_2$	022	$N_1P_2K_2$	122	$N_2P_2K_2$	222

É complexa a utilização do delineamento em blocos ao acaso num ensaio fatorial 3^3 , que apresenta um total de 27 tratamentos. Nesses casos efetua-se o confundimento de dois graus de liberdade da interação tripla com blocos, resultando em três blocos incompletos, em que cada um contempla nove tratamentos.

A interação tripla no experimento fatorial 3^3 possui oito graus de liberdade, o que possibilita obter quatro grupos de confundimento, denominados por Yates (1937) de grupos Z, W, X e Y, em que cada um utiliza dois graus de liberdade da interação.

Para estruturar os blocos incompletos, usam-se as equações de geometria finita, que são fundamentadas no sistema de numeração ternário, utilizando os algarismos 0, 1 e 2.

Quando ocorrer o número 3, o mesmo corresponderá ao zero no sistema ternário, o 4 corresponderá ao 1, o cinco ao 2, o 6 ao zero (0), e assim sucessivamente.

Denominando os níveis dos fatores por x_1, x_2, x_3 , para gerar os quatro grupos de confundimento designados por Yates (1937), como Z, W, X e Y, considerar as seguintes equações da geometria finita:

Grupos	Equações da geometria finita
Z	$x_1 + x_2 + x_3 = 0, 1, 2$
W	$2x_1 + x_2 + x_3 = 0, 1, 2$
X	$x_1 + 2x_2 + x_3 = 0, 1, 2$
Y	$x_1 + x_2 + 2x_3 = 0, 1, 2$

A equação $x_1 + x_2 + x_3 = 0, 1, 2$ promove a obtenção do plano experimental do grupo Z, em que:

$$\begin{aligned} x_1 + x_2 + x_3 = 3 &\rightarrow 0 & x_1 + x_2 + x_3 = 4 &\rightarrow 1 \\ x_1 + x_2 + x_3 = 5 &\rightarrow 2 & x_1 + x_2 + x_3 = 6 &\rightarrow 0 \end{aligned}$$

Bloco 1		Bloco 2		Bloco 3	
$x_1 + x_2 + x_3 = 0$		$x_1 + x_2 + x_3 = 1$		$x_1 + x_2 + x_3 = 2$	
000	$0 + 0 + 0 = 0$	001	$0 + 0 + 1 = 1$	002	$0 + 0 + 2 = 2$
012	$0 + 1 + 2 = 0$	010	$0 + 1 + 0 = 1$	011	$0 + 1 + 1 = 2$
021	$0 + 2 + 1 = 0$	022	$0 + 2 + 2 = 1$	020	$0 + 2 + 0 = 2$
102	$1 + 0 + 2 = 0$	100	$1 + 0 + 0 = 1$	101	$1 + 0 + 1 = 2$
111	$1 + 1 + 1 = 0$	112	$1 + 1 + 2 = 1$	110	$1 + 1 + 0 = 2$
120	$1 + 2 + 0 = 0$	121	$1 + 2 + 1 = 1$	122	$1 + 2 + 2 = 2$
201	$2 + 0 + 1 = 0$	202	$2 + 0 + 2 = 1$	200	$2 + 0 + 0 = 2$
210	$2 + 1 + 0 = 0$	211	$2 + 1 + 1 = 1$	212	$2 + 1 + 2 = 2$
222	$2 + 2 + 2 = 0$	220	$2 + 2 + 0 = 1$	221	$2 + 2 + 1 = 2$

Aplicando o mesmo procedimento, são estruturados os grupos Z, W, X e Y. Os quatro grupos de confundimento do fatorial 3^3 são perfeitamente balanceados, pois qualquer fator ocorre três vezes em cada um dos níveis e em qualquer bloco incompleto.

No caso do fatorial 3^3 , o confundimento é parcial de 2 graus de liberdade da interação $N \times P \times K$. É parcial porque a interação possui 8 graus de liberdade, de modo que serão confundidos apenas 2 graus de liberdade e os demais restantes informam sobre a interação tripla (parte não confundida).

De acordo com Pimentel Gomes (2000), os quatro grupos Z, W, X e Y estão assim formulados:

Grupo Y			Grupo Z		
1º bloco	2º bloco	3º bloco	1º bloco	2º bloco	3º bloco
000	001	002	000	001	002
011	012	010	012	010	011
022	020	021	021	022	020
101	102	100	102	100	101
112	110	111	111	112	110
120	121	122	120	121	122
202	200	201	201	202	200
210	211	212	210	211	212
221	222	220	222	220	221
Grupo W			Grupo X		
1º bloco	2º bloco	3º bloco	1º bloco	2º bloco	3º bloco
000	001	002	000	001	002
012	010	011	011	012	010
021	022	020	022	020	021
101	102	100	102	100	101
110	111	112	110	111	112
122	120	121	121	122	120
202	200	201	201	202	200
211	212	210	212	210	211
220	221	222	220	221	222

ANOVA (tabela 6.4.2.1) do fatorial 3^3 com e sem confundimento com três repetições:

Tabela 6.4.2.1. ANOVA do fatorial 3^3 com e sem confundimento.

Sem confundimento		Com confundimento	
F.V.	G.L.	F.V.	g.l.
N	2	N	2
P	2	P	2
K	2	K	2
$N \times P$	4	$N \times P$	4
$N \times K$	4	$N \times K$	4
$P \times K$	4	$P \times K$	4
$N \times P \times K$	8	$N \times P \times K$ (parte não confundida)	$6 = 8 - 2$
(Tratamentos)	(26)	Tratamento (parte não confundida)	$24 = 26 - 2$
Blocos	2		8
Resíduo	52		48
Total	80		80

6.4.3. Confundimento do fatorial 3^2

Os experimentos fatoriais 3×3 ou 3^2 consistem na aplicação de dois fatores, cada um com três níveis, resultando num conjunto de nove tratamentos. No caso de um experimento com adubação NPK, em que cada fator apresenta níveis equidistantes e representados por 0, 1, 2:

N_0P_0	N_0P_1	N_0P_2	N_1P_0	N_1P_1	N_1P_2	N_2P_0	N_2P_1	N_2P_2
00	01	02	10	11	12	20	21	22

Equações da geometria finita para gerar os dois grupos (x_1, x_2 : níveis dos fatores).

Grupos	Equações da geometria finita
1	$x_1 + x_2 = 0, 1, 2$
2	$x_1 + 2x_2 = 0, 1, 2$

Grupo1					
Bloco1		Bloco2		Bloco3	
$x_1 + x_2 = 0$		$x_1 + x_2 = 1$		$x_1 + x_2 = 2$	
00	$0 + 0 = 0$	01	$0 + 1 = 1$	02	$0 + 2 = 2$
12	$1 + 2 = 0$	10	$1 + 0 = 1$	11	$1 + 1 = 2$
21	$2 + 1 = 0$	22	$2 + 2 = 1$	20	$2 + 0 = 2$

Grupo2					
Bloco1		Bloco2		Bloco3	
$x_1 + 2x_2 = 0$		$x_1 + 2x_2 = 1$		$x_1 + 2x_2 = 2$	
00	$0 + 2 \times 0 = 0$	02	$0 + 2 \times 2 = 1$	01	$0 + 2 \times 1 = 2$
11	$1 + 2 \times 1 = 0$	10	$1 + 2 \times 0 = 1$	12	$1 + 2 \times 2 = 2$
22	$2 + 2 \times 2 = 0$	21	$2 + 2 \times 1 = 1$	20	$2 + 2 \times 0 = 2$

Observando que:

$$\begin{aligned} x_1 + x_2 + x_3 = 3 &\rightarrow 0 & x_1 + x_2 + x_3 = 4 &\rightarrow 1 \\ x_1 + x_2 + x_3 = 5 &\rightarrow 2 & x_1 + x_2 + x_3 = 6 &\rightarrow 0 \end{aligned}$$

Os dois grupos de confundimento apresentados do fatorial 3^2 são perfeitamente balanceados, pois qualquer fator ocorre uma vez em cada um dos níveis, em qualquer bloco incompleto. A tabela 6.4.3.1 apresenta a ANOVA do fatorial 3^2 com e sem confundimento. Segue um experimento, considerando cinco repetições, com casualização em blocos de adubação NP de uma espécie florestal.

Tabela 6.4.3.1. ANOVA do fatorial 3^2 com e sem confundimento.

Sem confundimento		Com confundimento	
F.V.	G.L.	F.V.	G.L.
N	2	N	2
P	2	P	2
$N \times P$	4	$N \times P$ (parte não confundida)	2
(Tratamentos)	(8)	(Tratamento) (parte não confundida)	6 = 8-2
Blocos	4		14
Resíduo	32		24
Total	44		44

6.5. NOÇÕES SOBRE FATORIAIS FRACIONÁRIOS OU FRACIONADOS

A grande desvantagem dos experimentos fatoriais, quando o número de fatores é relativamente grande, é o aumento do número de combinações (C) e, conseqüentemente, um grande acréscimo no número de parcelas a serem implantadas. Por exemplo, nos ensaios fatoriais 2^k ocorrem 512 combinações quando o número de fatores é igual a nove. O número de combinações cresce exponencialmente mesmo testando dois níveis para cada fator, o que pode inviabilizar a realização do experimento.

k	2	3	4	5	6	7	8	9
2^k	4	8	16	32	64	128	256	512

Um planejamento com seis fatores (2^6) requer $C = 64$ combinações para serem testadas, isto é, com esses fatores, estima-se uma média, 6 efeitos principais, 15 efeitos de interação com 2 fatores, 20 efeitos de interação com 3 fatores, 15 efeitos de interação com 4 fatores, 6 efeitos de interação com 5 fatores e uma interação com 6 fatores. Por outro lado, os efeitos de mais alta ordem são normalmente não significativos e, geralmente, podem ser descartados, assim despertando o interesse prático na utilização de frações de planejamentos fatoriais.

Quando o ensaio é planejado com muitos fatores, os efeitos principais dos fatores, as interações duplas ou triplas, são geralmente as mais importantes. Dessa forma, é possível reduzir o número de combinações e usar o procedimento denominado de planejamento fatorial fracionário ou fracionado. Os fatoriais fracionados são usados para reduzir o número de combinações e são importantes para efetuar a blocagem das combinações.

Conagin e Jorge (1977) citam o fatorial fracionário $(\frac{1}{5})5^3$ que é constituído por apenas 25 das 125 combinações do fatorial completo, podendo, assim, ter os seguintes tratamentos que constituiriam um único bloco.

111	134	152	125	143
254	222	245	213	231
342	315	333	351	324
435	453	421	444	412
523	541	514	532	555

Andrade e Noleto (1985) apresentam os fatoriais $(\frac{1}{2})4^3$ e $(\frac{1}{4})4^4$. O esquema $(\frac{1}{2})4^3$ apresenta apenas 32 das 64 combinações do fatorial completo, enquanto o fatorial $(\frac{1}{4})4^4$ apresenta 64 das 256 combinações possíveis.

Fatorial $(\frac{1}{2})4^3$							
Bloco1				Bloco2			
000	101	202	303	003	102	201	300
011	110	213	312	012	120	210	311
022	123	220	321	021	113	223	322
033	132	231	330	030	131	232	333

Fatorial $(\frac{1}{4})4^4$							
Bloco1		Bloco2		Bloco3		Bloco4	
0000	2022	0032	2010	0013	2031	0021	2003
0131	2113	0103	2121	0122	2100	0110	2132
0212	2230	0220	2202	0201	2223	0233	2211
0323	2301	0311	2333	0330	2312	0302	2320
1011	3033	1023	3001	1002	3020	1030	3012
1120	3102	1112	3130	1133	3111	1101	3123
1203	3221	1231	3213	1210	3232	1222	3200
1332	3310	1300	3322	1321	3303	1313	3331

6.5.1. Construção da matriz de planejamento do experimento fatorial 2^k

Há uma forma sistemática de estruturar a matriz de planejamento do experimento. Acrescente-se a isso que a construção da matriz de planejamento de um fatorial 2^3 , em que ocorrem 3 fatores de controle em 2 níveis (codificados por +1 e -1), resulta em $2^3 = 8$ combinações (C).

- Para o fator de controle 1 (x_1), escrever na coluna x_1 , o número (1) com sinal alternado seguidamente: -1, +1, -1, +1, -1, +1, -1, +1;
- Para o fator de controle 2 (x_2), escrever na coluna x_2 , o número (1) com o sinal alternado em pares: -1, -1, +1, +1, -1, -1, +1, +1;
- Para o fator de controle 3 (x_3), escrever na coluna de x_3 , o número (1) com o sinal alternado em grupos de 4: -1, -1, -1, -1, +1, +1, +1, +1. Resultando em:

		Fatores		
C	I	x_1	x_2	x_3
1	1	-1	-1	-1
2	1	+1	-1	-1
3	1	-1	+1	-1
4	1	+1	+1	-1
5	1	-1	-1	+1
6	1	+1	-1	+1
7	1	-1	+1	+1
8	1	+1	+1	+1

Desse modo, generalizando para o fatorial 2^k , pode-se adotar o procedimento:

- Para x_1 : o sinal da unidade (1), na coluna de x_1 , é alternado seguidamente em grupos de $2^0 = 1$;
- Para x_2 : o sinal da unidade (1), na coluna x_2 , é alternado seguidamente em grupos de $2^1 = 2$, isto é, em pares;
- Para x_3 : o sinal da unidade (1), na coluna de x_3 , é alternado em grupos de $2^2 = 4$;
- Para x_4 : o sinal da unidade (1), na coluna de x_4 , é alternado em grupos de $2^3 = 8$, isto é, $2^{4-1} = 8$;
- Para x_k : o sinal da unidade (1), na coluna de x_k , é alternado em grupos de 2^{k-1} , inicialmente 2^{k-1} vezes com o número (-1) e, em seguida, 2^{k-1} vezes o número (+1).

Essa forma padrão de organizar a matriz de planejamento do experimento garante que, as colunas de $x_1, x_2, \dots, x_k$, sejam ortogonais entre si. A figura 6.5.1.1 ilustra geometricamente as combinações do fatorial 2^3 .

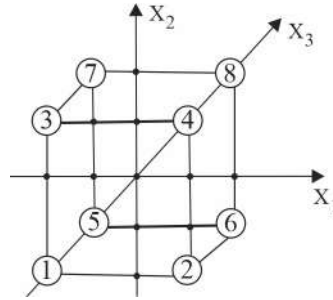


Figura 6.5.1.1. Representação geométrica do fatorial 2^3 . Fonte: Carpinette (2009).

No caso do fatorial fracionário 2^{k-p} , em que k é o número de fatores e 2^{k-p} combinações:

$$2^{k-p} = \frac{1}{2^p} \times 2^k$$

Para o teste de cinco fatores, considerando o fatorial completo, haveria a necessidade de estabelecer 32 combinações. Entretanto, usando o fatorial fracionário 2^{5-1} , o número de combinações seria reduzido para 16.

$$2^{k-p} = 2^{5-1} = \frac{1}{2^1} \times 2^5 = 16$$

No caso de aplicar o fatorial fracionário 2^{5-2} , o número de combinações passaria para 8.

$$2^{k-p} = 2^{5-2} = \frac{1}{2^2} \times 2^5 = 8$$

O uso do fatorial fracionário permite estabelecer menos combinações, diminuindo, desse modo, o custo da experimentação. Entretanto, a análise dos resultados pode ficar prejudicada, gerando perda de informação e de precisão nas estimativas dos efeitos. Por outro lado, os experimentos fracionários são recomendados para uso em pesquisas preliminares e exploratórias (*screening experiments*), principalmente quando o ensaio não demanda longo período para finalização, permitindo, desse jeito, o estudo e seleção dos fatores que podem ser mais importantes para atingir os objetivos da experimentação.

Geralmente os experimentos exploratórios são realizados nos estágios iniciais de um programa de pesquisa, quando é previsto que alguns fatores terão pouco ou nenhum efeito na resposta. Os fatores que forem identificados como importantes serão investigados mais profundamente em experimentos subsequentes.

A concepção básica do experimento fatorial fracionário é, justamente, estabelecer ensaios com muitos fatores em frações menores, visando estudar os efeitos mais importantes e assim planejar para o futuro experimentos fatoriais completos com menos fatores, obviamente utilizando os fatores que realmente são mais importantes.

Por outro lado, é possível implantar uma sequência de experimentos pequenos e eficientes, visando interagir as informações decorrentes desses vários ensaios, observando as vantagens do aprendizado, dentro do conceito da experimentação sequencial, que os fatoriais fracionários podem oferecer no desenvolvimento e na aplicação.

A ideia básica é dividir o experimento fatorial em frações, ou seja, definir os efeitos que serão propositalmente confundidos. O confundimento entre efeitos é o mecanismo empregado para reduzir o número de combinações, em que o objetivo é desprezar os fatores de pouca importância, permitindo o planejamento de ensaios com um menor número de fatores e estudando os efeitos com interações baixas.

Consideram-se as designações dos planejamentos, por exemplo, de fatoriais fracionários de frações metade 2^{k-1} , um quarto 2^{k-2} , um oitavo 2^{k-3} , em que k é o número de fatores (variáveis). Os termos no expoente (-1, -2, -3), por sua vez, indicam o grau de fracionamento, pois $2^{-1} = \frac{1}{2}$, $2^{-2} = \frac{1}{4}$ e $2^{-3} = \frac{1}{8}$. Tais termos indicam também o número de geradores necessários para planejar o experimento. Isto é, os ensaios com meia fração terão um gerador; os ensaios com fração um quarto terão dois geradores; e os projetos com um oitavo fracionado terão três geradores, e assim sucessivamente.

Na literatura existem tabelas que mostram os geradores para os projetos mais usados. No item 6.5.4, a tabela 6.5.4.1 mostra alguns exemplos de planejamentos. Em um experimento fatorial fracionário:

- i) Os k fatores são examinados e 2^{k-p} testes de combinações são realizados;
- ii) Exige que p variáveis sejam incorporadas no experimento 2^{k-p} , tal que colunas de interação da matriz de planejamento são usadas para a introdução dessas variáveis.

6.5.2. Planejamento de um fatorial 2^{k-p} e escolha das combinações a serem testadas.

a) Obtenção dos confundimentos de um fatorial fracionário $2^{k-p} = 2^{3-1} = 4$

Considerar, por exemplo, o fracionário $2^{k-p} = 2^{3-1} = 4$, em que a matriz do planejamento do fatorial 2^3 completo, tabela 6.5.2.1, é a seguinte:

Fatores: 1, 2, 3

Interações duplas: 12, 13, 23

Interação tripla: 123

Tabela 6.5.2.1. Matriz do planejamento do fatorial 2^3 completo.

C	I	1	2	3	12	13	23	123
1	1	-1	-1	-1	1	1	1	-1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1
6	1	1	-1	1	-1	1	-1	-1
7	1	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1	1

A partir da ordenação da interação tripla (123), a matriz ordenada para os valores de 1 e -1 (tabela 6.5.2.2) é:

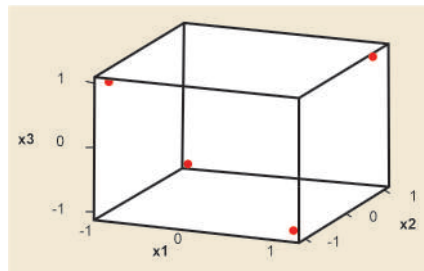
Tabela 6.5.2.2. Matriz do fatorial 2^3 com a interação 123 ordenada.

C	I	1	2	3	12	13	23	123
1	1	1	-1	-1	-1	-1	1	1
2	1	-1	1	-1	-1	1	-1	1
3	1	-1	-1	1	1	-1	-1	1
4	1	1	1	1	1	1	1	1
5	1	-1	-1	-1	1	1	1	-1
6	1	1	1	-1	1	-1	-1	-1
7	1	1	-1	1	-1	1	-1	-1
8	1	-1	1	1	-1	-1	1	-1

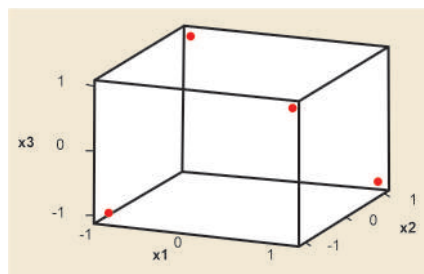
Observando o planejamento fatorial 2^3 , verifica-se que a tabela 6.5.2.2 apresenta, propositadamente, uma estruturação, tal que a coluna da interação tripla 123 está com os números ordenados na forma decrescente, isto é, (1, 1, 1, 1) e depois (-1, -1, -1, -1). Desse modo, está sendo realizada uma seleção de uma meia fração desse planejamento. Então se denomina o valor 12 como gerador dessa fração particular, em que o elemento identidade possui o mesmo sinal positivo para as quatro combinações ordenadas (1, 2, 3, 4), assim como se denomina $I = 123$ de relação de definição para esse planejamento fracionado. A fração com o sinal de identidade positivo na relação de definição é chamada de fração principal, enquanto a fração com sinal negativo na interação 123, isto é, $I = -123$, é denominada de fração alternada.

Representações gráficas das frações principal, alternada e fatorial completo 2^3 :

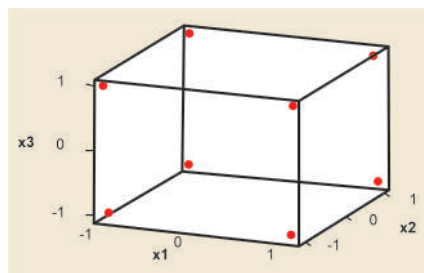
a1) Fração principal com relação definidora $I = 123$:



a2) Fração alternada com relação definidora $I = -123$:



a3) Fatorial completo:



Na tabela 6.5.2.3, os contrastes na meia fração selecionada que estimam o fator 1, $(1, -1, -1, 1)$, são os mesmos que estimam os contrastes da interação 23, assim como os contrastes que estimam o fator 2 são os mesmos que estimam a interação 13, e igualmente os contrastes do fator 3 são os mesmos que estimam a interação 12.

Tabela 6.5.2.3. Matriz considerando a meia fração principal do fatorial 2^3 .

C	I	1	2	3	12	13	23	123
1	1	1	-1	-1	-1	-1	1	1
2	1	-1	1	-1	-1	1	-1	1
3	1	-1	-1	1	1	-1	-1	1
4	1	1	1	1	1	1	1	1

Então, considerando essa meia fração, percebe-se que não é necessário estimar as interações 23, 13 e 12, haja vista que o efeito de cada uma dessas interações está confundido com um efeito principal. Assim sendo, as mesmas ficam fora do modelo, originando a matriz básica do planejamento $2^{k-p} = 2^{3-1} = 4$, em que, para planejar o fatorial 2^{3-1} , uma terceira variável x_3 (fator 3) precisa ser introduzida no planejamento, ou seja, a variável x_3 será a interação x_1x_2 , resultando no gerador $3 = 12$.

C	I	x_1	x_2	$x_3 = x_1x_2$
1	1	1	-1	-1
2	1	-1	1	-1
3	1	-1	-1	1
4	1	1	1	1

b) Obtenção dos confundimentos de um fatorial fracionário $2^{k-p} = 2^{4-1} = 8$.

Na tabela 6.5.2.4, observar as 16 combinações da matriz de planejamento do fatorial completo 2^4 , em que, considerando a interação 1234, verifica-se que os tratamentos de 1 a 8 apresentam o valor (+1), os quais serão testados no planejamento 2^{4-1} , conforme a tabela 6.5.2.5.

Tabela 6.5.2.4. Planejamento da matriz do fatorial completo 2^4 .

C	I	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
1	1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1
2	1	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1
3	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
4	1	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1
5	1	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1
6	1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1
7	1	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	-1
10	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1
11	1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	1	-1
12	1	1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	-1
13	1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1
14	1	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1
15	1	-1	1	1	1	-1	-1	-1	1	1	1	-1	-1	-1	1	-1
16	1	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	-1

Tabela 6.5.2.5. Meia fração principal do fatorial 2^4 . Valor (+1) na interação 1234.

C	I	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
1	1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1
2	1	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1
3	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
4	1	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1
5	1	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1
6	1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1
7	1	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Quando os coeficientes da interação 1234 são iguais a +1, nota-se que os contrastes da coluna 1 são iguais aos contrastes da coluna 234, isto é, $(1 = 234)$, assim como ocorrerão: $(2 = 134)$, $(3 = 124)$, $(4 = 123)$, $(12 = 34)$, $(13 = 24)$, $(14 = 23)$ e a média $(I) = 1234$, ou seja, esses são os efeitos confundidos do planejamento 2^{4-1} .

Por outro lado, para planejar o fatorial 2^{4-1} , uma quarta variável x_4 (fator 4) precisa ser introduzida no planejamento. O uso do contraste da interação 123 para representar o fator 4, o qual utilizará os sinais correspondentes ao efeito 123, resultará no plano do fatorial fracionário 2^{4-1} , em que o gerador será $4 = 123$.

	y	Fatores			
		x_1	x_2	x_3	$x_4 = x_1x_2x_3$
1	y_1	-1	-1	-1	-1
2	y_2	+1	-1	-1	+1
3	y_3	-1	+1	-1	+1
4	y_4	+1	+1	-1	-1
5	y_5	-1	-1	+1	+1
6	y_6	+1	-1	+1	-1
7	y_7	-1	+1	+1	-1
8	y_8	+1	+1	+1	+1

As combinações lineares de um planejamento fracionário são obtidas por:

$$L_i = \frac{\text{Contraste}_i}{N/2}$$

Os contrastes são calculados inserindo os valores da variável dependente y após o sinal correspondente a cada unidade experimental, em que $N = 2^{k-p}$.

Por exemplo, no caso do fator 1 (x_1) que está confundido com a interação 234:

$$1 + 234 = \frac{(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)}{4}$$

Denota-se que o efeito do fator 1 está confundido (tabela 6.5.2.5) com a interação. A desvantagem do fatorial fracionário é a impossibilidade de conhecer quais as proporções de contribuição do efeito do fator 1 e da interação entre os fatores 2, 3 e 4 para o valor da combinação linear obtida.

O item 6.5.3 apresenta outra forma de obter o padrão de confundimento de um fatorial fracionário, a qual é obtida a partir da relação definidora, em que o primeiro passo é definir o gerador. O gerador de um fatorial fracionário decorre da escolha que se faz para definir o nível dos fatores que estão sendo acrescentados. Considerar o exemplo do fatorial 2^{4-1} , quando se usa a coluna correspondente da interação $x_1x_2x_3$, para definir o quarto fator — isto é, $x_4 = x_1x_2x_3$ —, implica na criação do gerador $4 = 123$.

Quando se define que o gerador é igual a $4 = 123$, significa que houve referência às colunas dos (+1/-1) da matriz de planejamento, isto é, os valores da coluna 4 são iguais ao produto dos valores dos -1/+1 das colunas dos fatores 1, 2 e 3.

6.5.3. Gerador e relação definidora

Considerar o fatorial completo $2^k = 2^5 = 32$. Então, no caso do uso de um fatorial $2^{k-p} = 2^{5-2} = 8$, o qual corresponde a uma fração de $\frac{1}{4}$ do experimento 2^5 , em que somente oito testes (combinações) poderão ser realizados, o que corresponde, em um fatorial completo, a estabelecer apenas três variáveis (fatores). Surge, pois, a necessidade de introduzir duas variáveis no planejamento básico 2^3 , visando atender o objetivo de testar cinco fatores.

A matriz de planejamento, tabela 6.5.3.1, considerando os fatores 1, 2 e 3, adotando a interação 12 (ou seja, $4 = 12$) como fator 4 e interação 13 (isto é, $5 = 13$) como fator 5, resulta na matriz de planejamento do fatorial 2^{5-2} .

Tabela 6.5.3.1. Matriz de planejamento do fatorial 2^{5-2} .

C	I	1	2	3	4 = 12	5 = 13	23	123	y
1	1	-	-	-	+	+	+	-	y_1
2	1	+	-	-	-	-	+	+	y_2
3	1	-	+	-	-	+	-	+	y_3
4	1	+	+	-	+	-	-	-	y_4
5	1	-	-	+	+	-	-	+	y_5
6	1	+	-	+	-	+	-	-	y_6
7	1	-	+	+	-	-	+	-	y_7
8	1	+	+	+	+	+	+	+	y_8
	L_0	L_1	L_2	L_3	L_{12}	L_{13}	L_{23}	L_{123}	

$$L_i = \frac{\text{Contraste}_i}{N/2}$$

N : é o número de combinações

A matriz de planejamento do experimento $2^{k-p} = 2^{5-2} = 8$ é constituída de sete colunas e mais a coluna I (média). Assim, a média, os efeitos principais e as interações das cinco variáveis ou fatores estão confundidos em oito combinações lineares dos dados. O problema a ser equacionado é determinar quais efeitos estão confundidos entre si.

Por definição, as igualdades $4 = 12$ e $5 = 13$ são denominadas de geradores do experimento fatorial $2^{k-p} = 2^{5-2} = 8$. É fácil verificar que quando se multiplica uma coluna por ela mesma, há como resultado uma coluna com sinais positivos, que corresponde à coluna (I). Logo, como $4 = 12$ e $5 = 13$:

$$4 \times 4 = 12 \times 4 \quad 5 \times 5 = 13 \times 5$$

Resultando em:

$$I = 124 \text{ e } I = 135$$

Considerando que não se trata de uma operação algébrica, então:

$$124 = 135 \text{ ou } 124 \times 124 = 124 \times 135, \text{ ou seja, } I = 2345. \text{ Logo: } I = 124 = 135 = 2345.$$

A identidade $I = 124 = 135 = 2345$ é denominada como a relação definidora do planejamento fatorial $2^{k-p} = 2^{5-2} = 8$.

O padrão de confundimento é definido por meio da multiplicação de cada coluna da matriz de cálculo do experimento (tabela 6.5.3.1), no caso de $2^{k-p} = 2^{5-2} = 8$, pela relação definidora.

Logo, para coluna 1 (fator 1) no caso do fatorial 2^{5-2} :

$$1 \times I = 1 \times 124 = 1 \times 135 = 1 \times 2345 \rightarrow 1 = 24 = 35 = 12345$$

Significando que a função linear $L_1 = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)$, estima-se a soma dos efeitos $(1 + 24 + 35 + 12345)$.

Para a coluna 2 (fator 2), no caso do fatorial 2^{5-2} :

$$2 \times I = 2 \times 124 = 2 \times 135 = 2 \times 2345 \rightarrow 2 = 14 = 1235 = 345$$

Significando que a função linear $L_2 = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4 - y_5 - y_6 + y_7 + y_8)$, estima-se a soma dos efeitos $(2 + 14 + 1235 + 345)$.

Carpinetti (2009) apresenta os padrões de confundimento para as demais colunas:

L_3	estima $\rightarrow$	$3 + 1234 + 15 + 245$
L_{12}	estima $\rightarrow$	$12 + 4 + 235 + 1345$
L_{13}	estima $\rightarrow$	$13 + 234 + 5 + 1245$
L_{23}	estima $\rightarrow$	$23 + 134 + 125 + 45$
L_{123}	estima $\rightarrow$	$123 + 34 + 25 + 145$
L_0	estima $\rightarrow$	$média + \frac{1}{2}(124 + 135 + 2345)$

Carpinetti (2009) comenta que, se os efeitos predominantes são os efeitos principais e os de interação de dois fatores, o padrão de confundimento pode ser reduzido para:

L_0	$\rightarrow$	<i>média</i>	L_{12}	$\rightarrow$	$12 + 4$
L_1	$\rightarrow$	$1 + 24 + 35$	L_{13}	$\rightarrow$	$13 + 5$
L_2	$\rightarrow$	$2 + 14$	L_{23}	$\rightarrow$	$23 + 45$
L_3	$\rightarrow$	$3 + 15$	L_{123}	$\rightarrow$	$34 + 25$

O experimento 2^{5-2} apresenta, conforme mostra o quadro acima, um padrão de confundimento entre efeitos principais e de interação de dois fatores principais, o que dificulta a análise do ensaio. No caso de um plano fatorial 2^{5-1} — isto é, fração $\frac{1}{2}$ do experimento 2^5 , cujo planejamento básico é um fatorial 2^4 , em que a coluna da interação 1234 pode ser usada para representar a quinta variável —, resultará no gerador $5 = 1234$, e, como consequência, a relação definidora será $I = 12345$.

Por exemplo, considerando a coluna 1 (fator 1), no caso do fatorial 2^{5-1} :

$$1 \times I = 1 \times 12345 \rightarrow 1 = 2345$$

Calculando o padrão de confundimento para 2^{5-1} , conforme o procedimento usado no experimento 2^{5-2} , a matriz de cálculo do experimento básico (2^4) estimará, de acordo com Carpinetti (2009), os seguintes padrões de confundimento:

$L_0 = média + \frac{1}{2}(12345)$	$L_4 = 4 + 1235$	$L_{23} = 23 + 145$	$L_{124} = 124 + 35$
$L_1 = 1 + 2345$	$L_{12} = 12 + 345$	$L_{24} = 24 + 135$	$L_{134} = 134 + 25$
$L_2 = 2 + 1345$	$L_{13} = 13 + 245$	$L_{34} = 34 + 125$	$L_{234} = 234 + 15$
$L_3 = 3 + 1245$	$L_{14} = 14 + 235$	$L_{123} = 123 + 45$	$L_{1234} = 1234 + 5$

Considerando o exposto no quadro acima e assumindo que os efeitos de interação de 3, 4 e 5 fatores são desprezíveis, a matriz de cálculo do experimento básico (2^4) estimará os seguintes efeitos:

$L_0 \rightarrow \text{média}$	$L_4 \rightarrow 4$	$L_{23} \rightarrow 23$	$L_{124} \rightarrow 35$
$L_1 \rightarrow 1$	$L_{12} \rightarrow 12$	$L_{24} \rightarrow 24$	$L_{134} \rightarrow 25$
$L_2 \rightarrow 2$	$L_{13} \rightarrow 13$	$L_{34} \rightarrow 34$	$L_{234} \rightarrow 15$
$L_3 \rightarrow 3$	$L_{14} \rightarrow 14$	$L_{123} \rightarrow 45$	$L_{1234} \rightarrow 5$

O experimento 2^{5-1} , diferentemente do 2^{5-2} , não apresenta, conforme mostra o quadro acima, um padrão de confundimento entre efeitos principais e de interação entre dois fatores, o que facilita a análise do experimento.

Examinar os padrões de confundimentos, no caso do fatorial fracionário $2^{k-p} = 2^{3-1} = 4$, calculados pela multiplicação dos fatores pela relação definidora, em que $I = 123$. Além disso, observar, no item 6.5.2 (a), que os padrões de confundimentos foram obtidos de outra forma.

a) Multiplicação da relação definidora pelo fator 1:

$$1 \times I = 1 \times 123 = 23 \rightarrow 1 = 23$$

Destacando que não se trata de uma operação algébrica, quando se estima o efeito do fator 1, o efeito da interação 23 também é estimado.

b) Multiplicação da relação definidora pelo fator 2:

$$2 \times I = 2 \times 123 = 13 \rightarrow 2 = 13$$

Quando se estima o efeito do fator 2, o efeito da interação 13 também é estimado.

c) Multiplicação da relação definidora pelo fator 3:

$$3 \times I = 3 \times 123 = 12 \rightarrow 3 = 12$$

Quando se estima o efeito do fator 3, o efeito da interação 12 também é estimado.

Como visto no procedimento anterior, item 6.5.2 (a), só é necessário estimar os efeitos dos fatores 1, 2 e 3, pois as interações 12, 13 e 23 estão confundidas com esses efeitos. Esses efeitos que estão confundidos com os fatores são denominados de pares associados ou aliases. No caso de fatoriais fracionários 2^{k-1} , todos os efeitos possuem um par associado ou aliases. No caso presente do fatorial (2^{3-1}), os fatores possuem um par associado de segunda ordem.

Considerar os padrões de confundimentos, nos quais o caso do fatorial fracionário seja $2^{k-p} = 2^{4-1} = 8$, calculados pela multiplicação dos fatores pela relação definidora, em que $I = 1234$. Além disso, ver o item 6.5.2 (b), tabelas 6.5.2.4 e 6.5.2.5, em que os padrões de confundimentos foram obtidos de outra forma.

Para estabelecer o padrão de confundimento, considerando o fatorial 2^{4-1} , para o gerador $4 = 123$ com a relação definidora igual a $I = 1234$, primeiramente, multiplica-se a relação definidora $I = 1234$ pelos fatores e interações, informações explicitadas nas colunas da matriz de cálculo dos efeitos do planejamento básico. A tabela 6.5.3.2 apresenta a matriz de planejamento básico do fatorial 2^{4-1} .

	Fatores			Interações			
	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_4 = x_1x_2x_3$
1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	-1	-1	+1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	-1	+1	-1	-1	-1
5	-1	-1	+1	+1	-1	-1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	-1	+1	+1	-1	-1	+1	-1
8	+1	+1	+1	+1	+1	+1	+1

De acordo com a tabela 6.5.3.2, no caso do fatorial 2^{4-1} , com o gerador $4 = 123$ e a relação definidora $I = 1234$, há os seguintes padrões de confundimentos:

$$(1 = 234), (2 = 134), (3 = 124), (4 = 123), (12 = 34), (13 = 24), (23 = 14)$$

Quando se define que o gerador é igual a $4 = 123$, significa que se fez referência às colunas dos (+1/-1) da matriz de planejamento. Isto é, os valores da coluna 4 são iguais ao produto dos valores dos -1/+1 das colunas dos fatores 1, 2 e 3.

a) No caso da primeira coluna x_1 (fator 1), em que os contrastes da coluna x_1 são iguais aos contrastes da coluna 234:

$1 \times I = 1 \times 1234$ que implica $1 = 234$, então:

$$1 + 234 = \frac{(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)}{4}$$

$(1 + 234)$ corresponde à soma do efeito principal do fator 1 mais o efeito da interação entre os fatores 2, 3 e 4. Por outro lado, não é possível conhecer quais são as proporções entre esses dois efeitos, apenas se sabe que eles estão confundidos.

b) No caso da segunda x_2 coluna (fator 2), em que os contrastes da coluna x_2 são iguais aos contrastes da coluna 134:

$2 \times I = 2 \times 1234$ que implica $2 = 134$, ou seja:

$$2 + 134 = \frac{(-y_1 - y_2 + y_3 + y_4 - y_5 - y_6 + y_7 + y_8)}{4}$$

$(2 + 134)$ corresponde à soma do efeito principal do fator 2 mais o efeito da interação entre os fatores 1, 3 e 4. Contudo, não é possível conhecer quais são as proporções entre esses dois efeitos, apenas se denota que eles estão confundidos.

c) No caso da terceira coluna x_3 (fator 3), em que os contrastes da coluna x_3 são iguais aos contrastes da coluna 124:

$3 \times I = 3 \times 1234$ que implica $1 = 124$, ou seja:

$$3 + 124 = \frac{(-y_1 - y_2 - y_3 - y_4 + y_5 + y_6 + y_7 + y_8)}{4}$$

$(3 + 124)$ corresponde à soma do efeito principal do fator 3 mais o efeito da interação entre os fatores 1, 2 e 4. Entretanto, não se conhece quais são as proporções entre esses dois efeitos, apenas há indicação de que eles estão confundidos.

d) No caso da quarta coluna $x_1 x_2$ (interação entre os fatores 1 e 2), em que os contrastes da coluna $x_1 x_2$ são iguais aos contrastes da coluna 34:

$12 \times I = 12 \times 1234$ que implica $12 = 34$, ou seja:

$$3 + 124 = \frac{(-y_1 - y_2 - y_3 - y_4 + y_5 + y_6 + y_7 + y_8)}{4}$$

$(12 + 34)$ corresponde à soma do efeito da interação dos fatores 1 e 2 mais o efeito da interação entre os fatores 3 e 4. Entretanto, não se sabe quais são as proporções entre esses dois efeitos, apenas há indicação de que eles estão confundidos.

e) No caso da quarta coluna $x_1 x_3$ (interação entre os fatores 1 e 3), em que os contrastes da coluna $x_1 x_3$ são iguais aos contrastes da coluna 24:

$13 \times I = 13 \times 1234$ que implica $13 = 24$, ou seja:

$$13 + 24 = \frac{(+y_1 - y_2 + y_3 - y_4 - y_5 + y_6 - y_7 + y_8)}{4}$$

(13 + 24) corresponde à soma do efeito da interação entre os fatores 1 e 3 mais o efeito da interação entre os fatores 2 e 4. Por outro lado, não se sabe quais são as proporções entre esses dois efeitos, apenas há sinalização de que eles estão confundidos.

f) No caso da coluna x_2x_3 (interação entre os fatores 2 e 3), em que os contrastes da coluna x_2x_3 são iguais aos contrastes da coluna 14:

$23 \times I = 23 \times 1234$ que implica $23 = 14$, ou seja:

$$23 + 14 = \frac{(+y_1 + y_2 - y_3 - y_4 - y_5 - y_6 + y_7 + y_8)}{4}$$

(23 + 14) corresponde à soma do efeito da interação entre os fatores 2 e 3 mais o efeito da interação entre os fatores 1 e 4. Entretanto, não se sabe quais são as proporções entre esses dois efeitos, apenas há conhecimento que eles estão confundidos.

g) No caso da coluna $x_4 = x_1x_2x_3$ (interação entre os fatores 1, 2 e 3), em que os contrastes da coluna $x_1x_2x_3$ são iguais aos contrastes da coluna 4:

$123 \times I = 123 \times 1234$ que implica $123 = 4$, ou seja:

$$123 + 4 = \frac{(-y_1 + y_2 + y_3 - y_4 + y_5 - y_6 - y_7 + y_8)}{4}$$

(123 + 4) corresponde à soma do efeito da interação $x_1x_2x_3$ com o fator 4. Contudo, não se conhece quais são as proporções entre esses dois efeitos, apenas se sabe que eles estão confundidos.

Geralmente é assumido que o efeito principal possui um efeito preponderante, ou seja, a maior parte pertence ao efeito principal e a menor parte ao efeito da interação, principalmente em se tratando de uma interação da mais alta ordem, mas é apenas uma suposição baseada em uma regra geral.

Além das técnicas já citadas, com o objetivo de reduzir o número de pontos experimentais, foi criada a técnica que utiliza os delineamentos compostos centrais. Eles foram desenvolvidos inicialmente por Box e Wilson (1951), para estudo de funções polinomiais de resposta na indústria, em que o erro experimental, em geral, é reduzido e as condições do experimento são mais facilmente controláveis. Nessas condições é comum repetir apenas um tratamento, no caso, o relativo ao ponto central. O Delineamento Composto Central (DCC) é considerado um ótimo delineamento. O DCC, além de apresentar um número menor de tratamentos, em relação aos fatoriais completos, é importante na busca do ponto que mostre a resposta ótima.

O delineamento composto central é um delineamento, em alguns casos, simétrico e de segunda ordem, constituído de duas partes: o fatorial 2^k , com um ou mais pontos centrais e a parte axial. O capítulo 11 apresenta mais detalhes sobre esse delineamento.

Pimentel Gomes (2000) apresenta alguns tipos de delineamentos axiais de fatorial incompleto, como é o caso de um experimento com dois fatores (N e P) com os seguintes tratamentos, em que o ponto central (33) apresenta doses medianas de N e P :

		35		
		34		
13	23	33	43	53
		32		
		31		

O referido autor menciona também o delineamento axial FAO-ANDA com três níveis, para os fatores N , P e K , apresentando apenas sete tratamentos. Em alguns casos, podem ainda ser adicionadas as combinações 000 (testemunha/sem adubação) e (111 + calcário) para avaliar o efeito da calagem.

011	101	110
	111	
211	121	112

6.5.4. Definição de resolução em um experimento fracionário

Um fatorial fracionário possui resolução do tipo *III*, isto é, quando nenhum efeito principal está relacionado com outro efeito principal, mas pode estar relacionado com uma interação de segunda ordem, a qual, por sua vez, pode estar relacionada com interação de terceira ordem ou de ordem superior. Em resumo, significa que ao menos alguns efeitos principais serão confundidos com efeitos de interação de dois fatores.

O planejamento 2^{3-1} apresenta a resolução tipo *III*, em que o fator 1 está confundido com a interação 23, o fator 2 com a interação 13 e o fator 3 com a interação 12, assim como, $I = 123$. Em outras palavras, I possui três elementos. Assim sendo é escrito da seguinte forma: 2^{3-1}_{III} .

O planejamento 2^{5-2} também possui a resolução tipo *III*, pois, considerando os geradores $4 = 12$ e $5 = 13$ e a relação definidora $I = 124 = 135 = 2345$, o mesmo apresenta os seguintes padrões de confundimentos:

$1 = 24 = 35 = 12345$	$3 = 1234 = 15 = 245$	$13 = 234 = 5 = 1245$
$2 = 14 = 1235 = 345$	$12 = 4 = 235 = 1345$	$23 = 134 = 125 = 45$
$123 = 34 = 25 = 145$		

A resolução *III* é uma situação não recomendável, pois apresenta confundimento de um efeito principal de um fator com a interação com dois fatores. No caso do planejamento 2^{5-2} , há, por exemplo, o fator 4 confundido com o efeito da interação com os fatores 1 e 2, pois ($12 = 4 = 235 = 1345$). No caso do fatorial 2^{3-1} , há, por exemplo, o fator 1 confundido com a interação dos fatores 2 e 3.

A resolução do tipo IV estabelece que nenhum efeito principal esteja relacionado com outro efeito principal ou com interação de segunda ordem. Contudo estão confundidos com interação de terceira ordem, assim como apresentam interações de segunda ordem confundidas entre si. Um planejamento 2^{4-1} apresenta a resolução IV , pois $I = 1234$, em que I apresenta quatro elementos, sendo representado por 2_{IV}^{4-1} , tal que:

$$(1 = 234), (2 = 134), (3 = 124), (4 = 123), (12 = 34), (13 = 24), (23 = 14)$$

No exemplo do fatorial 2^{4-1} , com a relação definidora $I = 1234$, no caso do confundimento $4 = 123$, existe um efeito principal se confundindo com o efeito de interação de três fatores, resultando em uma resolução igual a IV , haja vista que:

Resolução igual a $1+3 = 4$, ou seja, um efeito principal (um fator) se confundindo com uma interação com três fatores.

No confundimento $23 = 14$, em que a interação de dois fatores se confunde com uma interação também com dois fatores ($2 + 2 = 4$), ocorre uma resolução igual a IV .

A tabela 6.5.4.1 apresenta alguns exemplos de planejamento de fatoriais fracionários e suas resoluções.

Tabela 6.5.4.1. Resoluções e geradores de fatoriais fracionários.

k	Resolução	Plano	Num. Comb.	Geradores
3	III	2_{III}^{3-1}	4	$(3 = \pm 12)$
4	IV	2_{IV}^{4-1}	8	$(4 = \pm 123)$
5	III	2_{III}^{5-2}	8	$(4 = \pm 12) (5 = \pm 13)$
5	V	2_{V}^{5-1}	16	$(5 = \pm 1234)$
6	III	2_{III}^{6-3}	8	$(4 = \pm 12) (5 = \pm 13) (6 = \pm 23)$
6	IV	2_{IV}^{6-2}	16	$(5 = \pm 123) (6 = \pm 234)$
6	VI	2_{VI}^{6-1}	32	$(6 = \pm 12345)$
7	III	2_{III}^{7-4}	8	$(4 = \pm 12) (5 = \pm 13) (6 = \pm 23) (7 = \pm 123)$
7	IV	2_{IV}^{7-3}	16	$(5 = \pm 123) (6 = \pm 234) (7 = \pm 134)$
7	V	2_{V}^{7-2}	32	$(6 = \pm 1234) (7 = 1245)$

A resolução do tipo V indica que nem os efeitos principais ou os efeitos das interações de segundo grau estão relacionados com algum efeito principal. Alguns dos efeitos principais estão confundidos com interação de quatro fatores, assim como as interações de segundo grau, ou seja, entre dois fatores, estão relacionadas entre si. Entretanto, as interações de segundo grau podem estar relacionadas com as interações de terceiro grau, tal que $I = 12345$, em que I apresenta cinco elementos. O

fatorial fracionário 2^{5-1} apresenta a resolução IV , sendo representado por 2^{5-1}_{IV} . Considerar, assim, os padrões de confundimentos, no caso do fatorial fracionário 2^{5-1} , com o gerador $5 = 1234$ e a relação definidora $1 = 12345$:

1 = 2345	4 = 1235	14 = 235	34 = 125	134 = 25
2 = 1345	12 = 345	23 = 145	123 = 45	234 = 15
3 = 1245	13 = 245	24 = 135	124 = 35	1234 = 5

Exercício 6.5.4.1. Considerar um experimento de adubação de uma espécie florestal, considerando um ensaio fatorial incompleto 2^{4-1}_{IV} com duas repetições — e que a tabela 6.5.4.2 apresenta as produções de madeira em m^3 por parcela para quatro fatores (1, 2, 3, 4) equidistantes. Nesse cenário, ajustar o modelo de segunda ordem.

Tabela 6.5.4.2. Dados (simulados) para a análise de um fatorial 2^{4-1}_{IV}

TRAT	Fatores				Respostas	
	1	2	3	4	Rep1	Rep2
1	-1	-1	-1	-1	43,0	44,0
2	1	-1	-1	1	95,6	98,5
3	-1	1	-1	1	43,0	48,7
4	1	1	-1	-1	62,1	67,3
5	-1	-1	1	1	71,8	75,4
6	1	-1	1	-1	57,4	50,6
7	-1	1	1	-1	76,5	68,5
8	1	1	1	1	91,8	88,7

Considerar os padrões de confundimentos, no caso do fatorial fracionário 2^{4-1}_{IV} , com o gerador $4 = 123$ e a relação definidora $1 = 1234$:

Relação de definição = 1234						
1 = 234	2 = 134	3 = 124	4 = 123	12 = 34	13 = 24	23 = 14

No modelo, colocam-se os efeitos de primeira ordem ($1 = 234$, $2 = 134$, $3 = 124$, $4 = 123$) e os de segunda ordem ($12 = 34$, $13 = 24$, $23 = 14$). Não devem ser colocados todos os componentes, pois a colocação impossibilitaria a realização da análise de variância, ou seja, é necessário optar, por exemplo, entre as interações 23 ou 14, pois 23 e 14 são altamente correlacionadas devido ao confundimento. Assim sendo, efeitos superpostos não devem ser colocados.

Seja o modelo recomendado para a análise:

$$Y_i = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + e_i$$

A tabela 6.5.4.3 apresenta os resultados da análise de regressão obtidos com o programa **Prog21**, em anexo no apêndice B.

$$\hat{Y} = 67,68125 + 8,81875x_1 + 0,64375x_2 + 4,90625x_3 + 9,00625x_4 \\ + 0,33125x_1x_2 - 9,28125x_1x_3 + 8,14375x_2x_3$$

Os resultados (tabela 6.5.34.3) mostram que os efeitos do fator 2 (x_2) e da interação entre o fator 1 e o fator 2 (x_{12}) são não significativos.

No capítulo 11, que trata sobre metodologias de superfícies de resposta, serão apresentadas várias análises sobre como obter a dose economicamente aconselhável em ensaios de adubação.

Tabela 6.5.4.3. Análise do modelo de um fatorial fracionário 2^{4-1}_{III} .

Variáveis	G.L.	Parâmetros Estimados	Erro padrão	Valor “t”	Pr > t
Intercepto	1	67,68125	0,88774	76,24	< 0,0001
x1	1	8,81875	0,88774	9,93	< 0,0001
x2	1	0,64375	0,88774	0,73	0,4890
x3	1	4,90625	0,88774	5,53	0,0006
x4	1	9,00625	0,88774	10,15	< 0,0001
x12	1	0,33125	0,88774	0,37	0,7187
x13	1	-9,28125	0,88774	-10,45	< 0,0001
x23	1	8,14375	0,88774	9,17	< 0,0001

6.6. EXPERIMENTOS FATORIAIS COM TRATAMENTOS ADICIONAIS

Na experimentação em ciências agrárias, em que o objetivo é obter informações complementares, visando diminuir o número de tratamentos, é comum incluir em ensaios com estrutura unifatorial ou fatorial, um ou mais tratamentos adicionais.

Pimentel Gomes (2000) afirma que, nos experimentos fatoriais, especialmente nos de $3 \times 3 \times 3$ com confundimento de dois graus de liberdade da interação tripla, a inclusão de tratamentos adicionais tem apresentado uso generalizado na experimentação em ciências agrárias, citando que o uso de tratamentos adicionais, nesses tipos de experimentos, não origina complicações nas análises estatísticas.

Esse autor apresenta, como exemplo, um ensaio de adubação de milho, em um esquema fatorial $3 \times 3 \times 3$, com casualização em três blocos, com nitrogênio (N), fósforo (P) e potássio (K), com confundimento de 2 graus de liberdade da interação tripla (grupo W). Em cada um dos três blocos de 9 parcelas foram incluídos cinco tratamentos adicionais: 000, 111, 111 + calcário (C), 111+micro-nutrientes (M) e 111 + micronutrientes + calcário. Além disso, cada bloco apresenta 14 parcelas, sendo nove referentes ao experimento fatorial e cinco inerentes aos tratamentos adicionais.

A análise estatística foi realizada separadamente para o experimento fatorial e para os tratamentos adicionais. Posteriormente, foi feita uma análise combinada envolvendo os tratamentos do experimento fatorial com os tratamentos adicionais. Na análise conjunta, considerando similaridade entre os dois quadrados médios dos resíduos, o número de graus de liberdade para o resíduo comum é a combinação (soma) dos graus de liberdade das duas análises anteriores. Por fim, foi calculado o quadrado médio do resíduo comum:

$$QM_{RES\ COMUM} = \frac{SQR_{FAT} + SQR_{AD}}{n_{FAT} + n_{AD}}$$

SQR_{FAT} : é a soma de quadrados do resíduo dos tratamentos fatoriais e n_{FAT} é o número de graus de liberdade da análise correspondente aos tratamentos fatoriais;

SQR_{AD} : é a soma de quadrados do resíduo dos tratamentos adicionais e n_{AD} é o número de graus de liberdade da análise referente aos tratamentos adicionais.

A estrutura experimental permite realizar duas análises de variâncias separadamente, sendo uma para os tratamentos comuns (fatorial) e outra para os adicionais. Portanto, como os tratamentos adicionais são aleatorizados conjuntamente com os comuns, as duas estimativas dos resíduos são comparáveis e podem ser reunidas para formação de um resíduo médio. Esse resíduo médio pode ser usado para estimar a precisão do experimento e testar alguns contrastes entre médias de tratamentos.

Yassin (2001) constatou, na resolução de alguns exemplos numéricos, que a inclusão de tratamentos adicionais, apesar de ser uma prática simples e comum, pode trazer dificuldades na interpretação dos testes de hipóteses. O autor recomenda que as análises sejam feitas matricialmente nesses casos, haja vista a impossibilidade de incluir, no modelo linear do esquema fatorial, os termos referentes aos tratamentos adicionais.

Yassin (2001) cita alguns problemas que foram encontrados com a inclusão de tratamentos adicionais nas análises dos exemplos analisados em sua pesquisa. Entre eles, foram detectados:

- a) Caso de não existência de ortogonalidade de efeitos entre todas as comparações de interesse, logo existem contrastes ambíguos. Por exemplo, o emprego de tratamento adicional mais de uma vez nos contrastes;
- b) Ocorrência de comparações confusas, nos casos de uso de mais de um tratamento adicional, em que se pretende comparar, por exemplo, a média do fatorial com as médias dos adicionais, pois essas médias envolvem vários tratamentos, não se sabendo em qual está a diferença detectada ou não;
- c) Ocorrência de dois experimentos em um só contraste, pelo fato de os tratamentos adicionais comporem outro conjunto de tratamentos.
- d) A utilização da testemunha mais de uma vez (repetição dos dados), fazendo com que a testemunha seja comparada novamente, em função de repetição de dados do tratamento adicional;
- e) As análises e as conclusões dos resultados, devido às ocorrências citadas anteriormente, nem sempre são fáceis de serem efetuadas.

Yassin (2001) efetuou as análises, por meio do software SAS, utilizando os procedimentos matriciais do PROC IML, usando matrizes completas e matrizes com restrição (Σ). Os exemplos apresentados e analisados foram os seguintes:

- a) Fatorial $(2 \times 2) + 3$, em um delineamento inteiramente casualizado balanceado;
- b) Fatorial $(3 \times 4) + 4$, em um delineamento com casualização em blocos;
- c) Fatorial $(3 \times 3) + 4$, em um delineamento inteiramente casualizado, sendo os dados desbalanceados referentes ao fatorial e balanceados referentes aos tratamentos adicionais.

Exercício 6.6.1. Considerar um experimento fatorial duplo casualizado em bloco com dois tratamentos adicionais $(3 \times 3 + 2)$, em que três tipos de adubação (A1, A2 e A3) foram considerados como fator A e três novos métodos de manejo (M1, M2 e M3) de uma determinada espécie florestal como fator M . O primeiro tratamento adicional T1 corresponde à testemunha, isto é, sem adubação e sem manejo, enquanto o tratamento adicional T2 não foi adubado, mas submetido ao método de manejo convencional, ou seja, usado normalmente pela empresa, mas diferente dos três métodos (M1, M2 e M3). A tabela 6.6.1 apresenta os volumes de madeira correspondentes aos dados (simulados), em $\text{m}^3 \text{ha}^{-1}$, medidos na data do final da rotação ideal.

Tabela 6.6.1. Volume de madeira $\text{m}^3 \text{ha}^{-1}$.

Tratamentos		Blocos			Total
		1	2	3	
A1	M1	157,5	168,0	126,0	451,5
A2	M1	210,0	220,5	231,0	661,5
A3	M1	105,0	126,0	157,5	388,5
A1	M2	189,0	105,0	115,5	409,5
A2	M2	262,5	231,0	283,5	777,0
A3	M2	105,0	126,0	115,5	346,5
A1	M3	147,0	105,0	168,0	420,0
A2	M3	157,5	168,0	147,0	472,5
A3	M3	157,5	126,0	168,0	451,5
T1		105,0	126,0	136,5	367,5
T2		138,5	137,5	159,0	435,0
Total		1.734,5	1.639,0	1.807,5	5.181,0

- a) Análise com casualização em blocos, considerando os 11 tratamentos:

$$C = \frac{5.181,0^2}{33} = 813.417,0$$

$$SQ_{TOTAL} = 157,5^2 + 168^2 + \dots + 137,5^2 + 159,0^2 - C = 69.551,0$$

$$SQ_{BLOCOS} = \frac{1}{11} (1.734,5^2 + 1.639,0^2 + 1.807,5^2) - C = 1.298,2273$$

$$SQ_{TRAT} = \frac{1}{3} (451,5^2 + 661,5^2 + \dots + 367,5^2 + 435,0^2) - C = 57.129,0$$

Tabela 6.6.2. ANOVA considerando todos os tratamentos.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	2	1.298,2273	649,1136	1,17
Tratamentos	10	57.129,0000	5.712,9000	10,27
Resíduo	20	11.123,7727	556,1886	
Total	32	69.551,0000		

b) Somas de quadrados dos fatores A (Adubação), M (Método de manejo) e da interação $A \times M$:

	M1	M2	M3	Total
A1	451,5	409,5	420,0	1.281,0
A2	661,5	777,0	472,5	1.911,0
A3	388,5	346,5	451,5	1.186,5
Total	1.501,5	1.533,0	1.344,0	4.378,5

$$SQ_A = \frac{1}{9}(1.281,0^2 + 1.911,0^2 + 1.186,5^2) - \frac{4.378,5^2}{27} = 34.471,5$$

$$SQ_M = \frac{1}{9}(1.501,5^2 + 1.533,0^2 + 1.344,0^2) - \frac{4.378,5^2}{27} = 2.278,5$$

$$SQ_{FATORIAL} = \frac{1}{9}(451,5^2 + 409,5^2 + \dots + 346,5^2 + 451,5^2) - \frac{4.378,5^2}{27} = 52.405,5$$

$$SQ_{A \times M} = SQ_{FATORIAL} - SQ_A - SQ_M = 52.405,5 - 34.471,5 - 2.278,5 = 15.655,5$$

c) Soma de quadrados dos tratamentos adicionais T1 e T2 (testemunhas):

T1	105,0	126,0	136,5	367,5
T2	138,5	137,5	159,0	435,0
Total				802,5

$$SQ_{TESTEMUNHA} = \frac{1}{3}(367,5^2 + 435,0^2) - \frac{802,5^2}{6} = 759,375$$

d) Soma de quadrados do contraste entre a média dos tratamentos fatoriais versus a média dos tratamentos adicionais T1 e T2:

$$SQ_{TESTEMUNHA \text{ VS } FATORIAL} = \frac{4.378,5^2}{27} + \frac{802,5^2}{6} - \frac{5.181,0^2}{33} = 3.964,1250$$

$$SQ_{TODOS \text{ TRATAMENTOS}} - SQ_{FATORIAL} = SQ_{ADICIONAIS \text{ VS } FATORIAL}$$

Observando-se que:

$$SQ_{TODOS \text{ TRATAMENTOS}} - SQ_{FATORIAL} = 57.129,0 - 52.405,5 = 4.723,5$$

$$SQ_{TESTEMUNHA \text{ VS } FATORIAL} = 4.723,5 - 759,3750 = 3.964,1250$$

e) Quadro completo da análise de variância:

Tabela 6.6.3. Quadro completo da ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	2	1.298,2273	649,1136	1,17
Tratamentos	(10)	57.129,0000	5.712,9000	10,27
Fatorial (Fat.)	8	52.405,5000	6.550,6875	11,78
<i>A</i> (Adubação)	2	34.471,5000	17.235,7500	30,99
<i>B</i> (Manejo)	2	2.278,5000	1.139,2500	2,05
<i>A</i> × <i>B</i>	4	15.655,5000	3.913,8750	7,04
Testemunhas (T1 e T2)	1	759,3750	759,3750	1,37 <i>ns</i>
Testemunhas vs Trat (Fat.)	1	3.964,1250	3.964,1250	7,13 **
Resíduo	20	11.123,7727	556,1886	
Total	32	69.551,0000		

$$F_{4; 20; 0,05} = 2,87 \quad F_{1; 20; 0,05} = 4,35$$

De acordo com a tabela 6.6.3, os tratamentos adicionais T1 e T2 são não significativos ($F = 1,37$), revelando que o método de manejo convencional não contribuiu para o aumento da produção da madeira. Por outro lado, existe significância entre o valor médio dos tratamentos fatoriais (combinações) com relação à média dos tratamentos adicionais ($F = 7,13$).

Considerando que a interação entre os fatores *A* e *M* é significativa, então há a necessidade de desdobrar os efeitos dos níveis do fator *A* dentro do fator *M* assim como dos efeitos dos níveis do fator *M* dentro do fator *A*.

f) Estudo de cada tipo de adubação (*A*) dentro dos métodos de manejo (*M*), ou seja, desdobramento dos graus de liberdade do fator manejo (*M*) mais o da interação *A*×*M*:

	M1	M2	M3	Total
A1	451,5	409,5	420,0	1.281,0
A2	661,5	777,0	472,5	1.911,0
A3	388,5	346,5	451,5	1.186,5
Total	1.501,5	1.533,0	1.344,0	4.378,5

$$SQ_{A1 D. M} = \frac{1}{3}(451,5^2 + 409,5^2 + 420,0^2) - \frac{1.281,0^2}{9} = 318,5$$

$$SQ_{A2 D. M} = \frac{1}{3}(661,5^2 + 777,0^2 + 472,5^2) - \frac{1.911,0^2}{9} = 15.753,5$$

$$SQ_{A3 D. M} = \frac{1}{3}(388,5^2 + 346,5^2 + 451,5^2) - \frac{1.186,5^2}{9} = 1.862,0$$

$$SQ_{A1 D. M} + SQ_{A2 D. M} + SQ_{A3 D. M} = SQ_M + SQ_{A \times M} = 2.278,5 + 15.655,5 = 17.934,0$$

F.V.	G.L.	S.Q.	Q.M.	F
A1 d. M	2	318,5000	159,2500	0,29 <i>ns</i>
A2 d. M	2	15.753,5000	7.876,7500	32,40 **
A3 d. M	2	1.862,0000	931,0000	1,67 <i>ns</i>
Resíduo	20	11.123,7727	556,1886	

Optou-se pelo teste de Tukey para comparação entre as médias fatoriais e pelo teste de Dunnett, para comparar as médias fatoriais com as duas testemunhas.

g) Teste de comparação múltipla de Tukey, considerando os tratamentos A2 d. M:

$$\Delta = q_{3; 20; 0,05} \sqrt{\frac{QM_{RES}}{J}} = 3,58 \sqrt{\frac{556,1886}{3}} = 48,74$$

A2M2	777,0	a
A2M1	661,5	b
A2M3	472,5	c

Os métodos de manejo diferem entre si, para o nível de significância $\alpha = 0,05$, pelo teste de Tukey, quando aplicado o tipo de adubo A2. O método de manejo M2 apresentou o melhor resultado quando usado o adubo A2.

h) Estudo de cada método de manejo dentro de cada tipo de adubação, o que corresponde a desdobrar os graus de liberdade do fator adubação mais o da interação Adubação×Método de manejo:

	M1	M2	M3	Total
A1	451,5	409,5	420,0	1.281,0
A2	661,5	777,0	472,5	1.911,0
A3	388,5	346,5	451,5	1.186,5
Total	1.501,5	1.533,0	1.344,0	4.378,5

$$SQ_{M1 D. A} = \frac{1}{3}(451,5^2 + 661,5^2 + 388,5^2) - \frac{1.501,5^2}{9} = 13.622,0$$

$$SQ_{M2 D. A} = \frac{1}{3}(409,5^2 + 777^2 + 346,5^2) - \frac{1.533,0^2}{9} = 36.039,5$$

$$SQ_{M3 D. A} = \frac{1}{3}(420,0^2 + 472,5^2 + 451,5^2) - \frac{1.344,0^2}{9} = 465,5$$

$$SQ_{M1 D. A} + SQ_{M2 D. A} + SQ_{M3 D. A} = SQ_A + SQ_{A \times M} = 34.471,5 + 15.655,5 = 50.127,0$$

F.V.	G.L.	S.Q.	Q.M.	F
M1 d. A	2	13.622,0000	6.811,0000	12,25
M2 d. A	2	36.039,5000	18.019,7500	32,40
M3 d. A	2	465,5000	232,7500	0,43
Resíduo	20	11.123,7727	556,1886	

Dados os três tipos de adubação, o método M3 apresenta o mesmo comportamento.

i) Teste de comparação múltipla de Tukey, considerando M1 d. A e M2 d. A:

$$\Delta = q_{3; 20; 0,05} \sqrt{\frac{QM_{RES}}{J}} = 3,58 \sqrt{\frac{556,1886}{3}} = 48,74$$

A2M1	661,5	a	A2M2	777,0	a
A1M1	451,5	b	A1M2	409,5	b
A3M1	388,5	c	A3M2	346,5	c

Os métodos de manejo M1 e M2 diferem quando submetidos aos três tipos de adubação, para o nível de significância $\alpha = 0,05$, pelo teste de Tukey. Os métodos M1 e M2 apresentam o melhor desempenho quando aplicado o adubo A2.

j) Teste de comparação múltipla de Dunnett para testar os tratamentos fatoriais (nove combinações) versus os tratamentos adicionais testemunhas (T1) e (T2):

No caso do teste de Dunnett, considerando que as médias dos níveis dos fatores foram obtidas a partir três repetições, assim como foram utilizadas três repetições para calcular as médias dos tratamentos adicionais, no cálculo da diferença mínima de significância (Δ), haverá a seguinte formulação:

$$\Delta = D_{9; 20; 0,05} \sqrt{\left(\frac{1}{3} + \frac{1}{3}\right) QM_{RES}} = 2,95 \sqrt{\frac{2 \times 556,1886}{3}} = 56,8051$$

Considerando as testemunhas T_1 e T_2 :

	Bloco1	Bloco2	Bloco3	Totais
T1	105,0	126,0	136,5	367,5
T2	138,5	137,5	159,0	435,0

Sendo as médias para as testemunhas T_1 e T_2 :

$$\bar{T}_1 = \frac{367,5}{3} = 122,5$$

$$\bar{T}_2 = \frac{435,0}{3} = 145,0$$

Considerando a testemunha T_1 ($\bar{T}_1 = 122,5$):

Combinações	Média	Média -122,5	
A2M2	259,0	136,5	*
A2M1	220,5	98,0	*
A2M3	157,5	35,0	<i>ns</i>
A1M1	150,5	28,0	<i>ns</i>
A3M3	150,5	28,0	<i>ns</i>
A1M3	140,0	17,5	<i>ns</i>
A1M2	136,5	14,0	<i>ns</i>
A3M1	129,5	7,0	<i>ns</i>
A3M2	115,5	-7,0	<i>ns</i>

As combinações A2M2 e A2M1 apresentam, para o nível de significância $\alpha = 0,05$, pelo teste de Dunnett, resultados significativos quando comparadas com a testemunha T_1 .

Considerando a testemunha T_2 ($\bar{T}_2 = 145,0$):

	Média	Média -145,0	
A2M2	259,0	114,0	*
A2M1	220,5	75,5	*
A2M3	157,5	12,5	<i>ns</i>
A1M1	150,5	5,5	<i>ns</i>
A3M3	150,5	5,5	<i>ns</i>
A1M3	140,0	-5,0	<i>ns</i>
A1M2	136,5	-8,5	<i>ns</i>
A3M1	129,5	-15,5	<i>ns</i>
A3M2	115,5	-29,5	<i>ns</i>

As combinações A2M2 e A2M1 apresentam, para o nível de significância $\alpha = 0,05$, pelo teste de Dunnett, resultados significativos quando comparadas com a testemunha T_2 .

Por outro lado, se a interação $A \times B$ tivesse sido não significativa, o que não ocorreu nesse exemplo, então as comparações seriam realizadas com as médias dos fatores diretamente, ou seja, não seriam realizados os desdobramentos de um fator dentro de outro.

6.7. BIBLIOGRAFIA

ALBUQUERQUE, A. M. M. O. **Análise estatística de um experimento aninhado com uma aplicação em dados de algodão**. Departamento de Estatística do Centro de Ciências e Tecnologia da UEPB. Campina Grande-PB, 2016. 36 p.

- ANDRADE, D. F.; NOLETO, A. Q. Exemplos de fatoriais fracionados $(\frac{1}{2})4^3$ e $(\frac{1}{4})4^4$, para o ajuste de modelos polinomiais quadráticos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 21, p. 677-680, 1986.
- BANZATTO, D. A.; KRONKA, N. **Experimentação agrícola**. Jaboticabal: FUNEP, 1989. 247 p.
- BOX, G.E.P.; WILSON, K.B. On the experimental attainment of optimum conditions. **J. Royal Stat. Soc.**, Oxford, v.13, p.1-45, 1951.
- CARPINETTI, L. C. R. **Planejamento e análise de experimentos**. Escola de Engenharia de São Carlos. Departamento de Engenharia de Produção. USP, 2009. 222 p.
- CONAGIN, A.; JORGE, J. P. N. Delineamentos $(\frac{1}{5})5^3$. **Bragantia**, 36: 23-58, 1977.
- PIMENTEL GOMES, F. **Curso de estatística experimental**. 14ª ed. São Paulo: Nobel, 2000. 475 p.
- SIMÕES, J. W. Métodos de produção de mudas de Eucalyptus. **IPEF**, n.1, p. 101-116, 1970.
- YASSIN, N. **Análise de experimentos fatoriais de dois fatores com tratamentos adicionais. Dissertação de mestrado**. Programa de pós-graduação em agronomia, área de concentração em estatística e experimentação agropecuária. UFLA. Lavras – MG, 2001. 171 p.
- YATES, F. **The design and analysis of factorial experiments**. Imperial Bureau of soil Science, Harpenden, Inglaterra, 1937.

CAPÍTULO 7

ANÁLISE DE GRUPOS DE EXPERIMENTOS

Este capítulo traz definições, conceitos, aplicações e exemplos ilustrativos sobre a análise conjunta de grupos de experimentos. Ele também trata das condições ideais para a sua implementação, assim como apresenta as recomendações para o estabelecimento de subgrupos, baseadas na ordem de grandeza dos quadrados médios residuais das análises individuais. Exibe, para os ensaios inteiramente ao acaso, os quadros das análises de variância com seus respectivos componentes para os modelos fixos, aleatórios e mistos. Mostra exemplos de aplicações, em um ensaio casualizado em blocos, considerando a interação significativa entre os tratamentos e locais, desdobrando os graus de liberdade de tratamentos mais o grau da interação. Por fim, aplica o teste de Tukey para as comparações dos tratamentos dentro dos locais.

7.1. CONSIDERAÇÕES GERAIS

A análise de grupos de experimentos é fundamental quando o objetivo é obter informações para toda uma região, composta de distintos sítios ou ecossistemas. Na experimentação florestal, é importante a instalação de grupos de ensaios que apresentem a mesma estrutura experimental, considerando o tempo, sítios, tipos florestais, variações de solo, entre outros. Dificilmente é possível conduzir um experimento com o objetivo de aplicar seus resultados para uma ampla região, haja vista que as condições climáticas e tipos de solos variam de um local para outro, interferindo, assim, em seus resultados. Apenas uma experimentação adequada pode promover uma generalização dos resultados.

Muitas vezes as condições climáticas e tipos de solos variam de um local para outro, o que pode afetar os resultados, principalmente quando nos experimentos estão envolvidas espécies vegetais. Dessa maneira, somente uma experimentação adequada pode dar uma solução definitiva para promover uma generalização dos resultados para uma ampla região, o que pode ser viabilizado pela análise conjunta de experimentos, permitindo estabelecer conclusões mais gerais e confiáveis.

Em uma pesquisa — cujo objetivo é estudar a viabilidade da introdução de uma espécie florestal exótica na Amazônia brasileira, onde as condições climáticas e variações pedológicas exigem implantar diversos ensaios distribuídos em locais selecionados e representativos, tal que os resultados possam ser utilizados com o sucesso esperado —, o planejamento para uma análise de grupos de experimentos está alicerçado no interesse de se obter resultados mais gerais. Para isso, é necessário haver a multiplicação dos locais e dos anos de experimentação, em que os ensaios deverão ser instalados da maneira mais simples possível e com custos não excessivos.

Tais ensaios possibilitam tirar conclusões sobre a natureza do ambiente em que foram instalados. Em grupos de experimentos, é recomendável estabelecer a estrutura dos ensaios individuais da maneira mais simples possível, mas sempre procurando atender aos objetivos do programa de pesquisa. O ideal é usar, se possível, os delineamentos inteiramente casualizados, em blocos ou quadrado latino. Os ensaios individualizados devem apresentar a mesma estrutura experimental, em outras palavras, os mesmos tratamentos e, se possível, com o mesmo número de repetições, o que facilitará realizar a análise de variância conjunta.

7.2. ANÁLISE ESTATÍSTICA DE GRUPOS DE EXPERIMENTOS

Box (1954) comenta, referindo-se às grandezas dos quadrados médios dos resíduos para grupos de experimentos, que quando a relação entre o maior e o menor quadrado médio do resíduo não for maior que quatro, sem restrições na análise conjunta, todos os ensaios poderão ser incluídos. Caso isso não aconteça, devem ser organizados subgrupos com quadrado médio de resíduos homogêneos, sendo as análises conjuntas feitas para cada subgrupo.

Para efetuar a análise conjunta, é necessário que ocorra homoscedasticidade, ou seja, homogeneidade de variâncias, o que na prática significa que os quadrados médios residuais não diferem muito entre si. A verificação da homogeneidade de variâncias pode ser feita, por exemplo, pelo teste do “ F_{Max} ” ou H' de Hartley ou pelo teste de Bartlett.

No caso de a hipótese de nulidade ser rejeitada, isto é, quando não há existência de homogeneidade de variâncias, pode-se proceder de duas formas distintas:

- Separar em subgrupos os experimentos com quadrados médios residuais semelhantes ou não muito discrepantes e fazer a análise conjunta;
- Proceder à análise de variância conjunta, utilizando o método proposto por Cochran (1954) que consiste, em função da heterocedasticidade, em ajustar os graus de liberdade do resíduo médio e os graus de liberdade da interação entre tratamentos e locais. Nesses casos, efetua-se a análise de variância conjunta como no caso do quadrado médio do resíduo homogêneo, mas com o número de graus de liberdade ajustado para consulta de tabelas F e de teste de comparação múltipla, em que n_1^* será usado para o resíduo médio e n_2^* para os graus de liberdade da interação locais $\times$ tratamentos, os quais são obtidos da seguinte maneira:

$$n_1^* = \frac{[QM_{RES(1)} + QM_{RES(2)} + \dots + QM_{RES(K)}]^2}{\frac{QM_{RES(1)}^2}{glRES(1)} + \frac{QM_{RES(2)}^2}{glRES(2)} + \dots + \frac{QM_{RES(K)}^2}{glRES(K)}}$$

$$glRES(K) < n_1^* < \sum_{k=1}^K glRES(k)$$

$$n_2^* = \frac{(I-1)(K-1)^2 V_1^2}{(K-2)V_2 + V_1^2}$$

$$V_1 = \frac{\sum_{k=1}^K QM_{RES(k)}}{K} \quad V_2 = \frac{\sum_{k=1}^K QM_{RES(k)}^2}{K}$$

I : é o número de tratamentos;

K : é o número de locais ou sítios;

$QM_{RES(k)}$: quadrado médio residual correspondente ao k -ésimo local ou experimento;

$glRES(k)$: número de graus de liberdade do resíduo inerente ao k -ésimo experimento.

De acordo com Banzatto e Kronka (1989), na análise de grupos de experimentos, a ordem de grandeza dos quadrados médios residuais dos ensaios individuais deve ser considerada em um mesmo grupo, quando os quadrados médios residuais não ultrapassem uma relação de 7:1.

As análises conjuntas podem ser realizadas, considerando as seguintes situações:

- Utilização, se possível, dos dados originais dos experimentos individuais;
- Utilização dos totais de tratamentos dos experimentos individuais;

c) Realização da análise por meio das médias dos ensaios individuais.

Considerar o exemplo de uma análise com casualização em cinco blocos, sendo oito ensaios individuais (locais) com seis tratamentos:

Ensaios individuais		Análise conjunta	
F.V.	G.L.	F.V.	G.L.
Blocos	4	Blocos d. Ensaio	32
Tratamentos	5	Tratamentos (T)	5
Resíduo	20	Locais (L)	7
Total	29	Interação T×L	35
		Resíduo médio	160
		Total	239

O número de graus de liberdade do resíduo médio é igual à soma dos graus de liberdade dos ensaios individuais, isto é, $8 \times 20 = 160$. De acordo com Banzatto e Kronka (1989) e Barbin (2019), na análise conjunta, em geral, o componente bloco dentro de experimentos não apresenta interesse prático.

7.2.1. Análise de variância individual (por experimento ou local)

Seja considerado o modelo matemático do ensaio casualizado em blocos:

$$x_{ij} = m + t_i + b_j + e_{ij}$$

Tabela 7.2.1.1. ANOVA com casualização em blocos (t_i : fixo e b_j : fixo).

F.V.	G.L.	E(Q.M.)
Blocos	$J - 1$	$\sigma^2 + I\phi_1$
Tratamentos	$I - 1$	$\sigma^2 + J\phi_2$
Resíduo	$(I - 1)(J - 1)$	σ^2
Total	$IJ - 1$	

$$\phi_1 = \frac{\sum_{j=1}^J b_j^2}{J - 1} \quad \phi_2 = \frac{\sum_{i=1}^I t_i^2}{I - 1}$$

x_{ij} : valor observado correspondente ao i -ésimo tratamento do j -ésimo bloco;

m : média geral;

t_i : efeito fixo do i -ésimo tratamento com $E(t_i) = t_i$, $E(t_i^2) = t_i^2$ e $\sum_{i=1}^I t_i = 0$;

b_j : efeito fixo do j -ésimo bloco com $E(b_j) = b_j$, $E(b_j^2) = b_j^2$ e $\sum_{j=1}^J b_j = 0$;

e_{ij} : erro experimental associado a x_{ij} , em que $e_{ij} \cap N(0, \sigma^2)$.

7.2.2. Análise de variância conjunta para o grupo

a) Componentes de variância de grupos de experimentos para o ensaio inteiramente ao acaso. Seja o modelo matemático:

$$x_{ijk} = m + t_i + l_k + (tl)_{ik} + e_{ijk}$$

$i = 1, \dots, I$ (tratamentos)

$j = 1, \dots, J$ (repetições)

$k = 1, \dots, K$ (locais ou experimentos)

x_{ijk} : valor observado correspondente ao i -ésimo tratamento, da j -ésima repetição, do k -ésimo local;

m : média geral;

t_i : efeito fixo referente ao i -ésimo tratamento;

l_k : efeito fixo referente ao k -ésimo local;

tl_{ik} : efeito aleatório da interação entre o i -ésimo tratamento e o k -ésimo local, que por definição é considerado aleatório, tal que $tl_{ik} \cap N(0, \sigma_{tl}^2)$;

e_{ijk} : erro experimental associado a x_{ijk} , em que $e_{ijk} \cap N(0, \sigma^2)$ e independentes.

A tabela 7.2.2.1 apresenta, de acordo com o modelo matemático do ensaio inteiramente ao acaso, os componentes de variância considerando os efeitos de tratamento e de local fixos:

Tabela 7.2.2.1. Modelo fixo (tratamento fixo e local fixo).

F.V.	G.L.	Q.M.	E(Q.M.)
Tratamentos (T)	$I - 1$	QM_1	$\sigma^2 + J\sigma_{tl}^2 + JK\phi_1$
Locais (L)	$K - 1$	QM_2	$\sigma^2 + J\sigma_{tl}^2 + IJ\phi_2$
Interação T×L	$(I - 1)(K - 1)$	QM_3	$\sigma^2 + J\sigma_{tl}^2$
Resíduo	$IK(J - 1)$	QM_4	σ^2
Total	$IKJ - 1$		

$$\phi_1 = \frac{\sum_{i=1}^I t_i^2}{I - 1} \quad \phi_2 = \frac{\sum_{k=1}^K l_k^2}{K - 1}$$

Os testes F , para testar os efeitos de tratamentos e de locais, terão o quadrado médio da interação entre tratamentos e locais como denominador. Para testar a interação Tratamentos×Locais, o teste de F terá o quadrado médio do resíduo como denominador.

A tabela 7.2.2.2 apresenta, de acordo com o modelo do ensaio inteiramente ao acaso, os componentes de variância para o efeito de tratamento fixo e o efeito de local aleatório.

$(tl)_{ik}$: efeito aleatório da interação entre tratamentos e locais, em que $tl_{ik} \cap N(0, \sigma_{tl}^2)$;

l_k : efeito aleatório referente ao k -ésimo local, tal que $l_k \cap N(0, \sigma_l^2)$.

Tabela 7.2.2.2. Modelo misto (tratamento fixo e local aleatório).

F.V.	G.L.	Q.M.	E(Q.M.)
Tratamentos (T)	$I - 1$	QM_1	$\sigma^2 + J\sigma_{tl}^2 + JK\phi_1$
Locais (L)	$K - 1$	QM_2	$\sigma^2 + J\sigma_{tl}^2 + IJ\sigma_l^2$
Interação T×L	$(I - 1)(K - 1)$	QM_3	$\sigma^2 + J\sigma_{tl}^2$
Resíduo	$IK(J - 1)$	QM_4	σ^2
Total	$IJK - 1$		

$$\phi_1 = \frac{\sum_{i=1}^I t_i^2}{I - 1}$$

Os testes F , para testar os efeitos de tratamentos e de locais, terão o quadrado médio da interação tratamentos e locais como denominador. Para testar a interação Tratamentos×Locais, o teste de F terá o quadrado médio do resíduo como denominador.

A tabela 7.2.2.3 apresenta, de acordo com o modelo matemático do ensaio inteiramente ao acaso, os componentes de variância considerando os efeitos de tratamentos e de local aleatórios.

Tabela 7.2.2.3. Modelo aleatório (tratamento aleatório e local aleatório).

F.V.	G.L.	Q.M.	E(Q.M.)
Tratamentos (T)	$I - 1$	QM_1	$\sigma^2 + J\sigma_{tl}^2 + JK\sigma_t^2$
Locais (L)	$K - 1$	QM_2	$\sigma^2 + J\sigma_{tl}^2 + IJ\sigma_l^2$
Interação T×L	$(I - 1)(K - 1)$	QM_3	$\sigma^2 + J\sigma_{tl}^2$
Resíduo	$IK(J - 1)$	QM_4	σ^2
Total	$IJK - 1$		

tl_{ik} : efeito aleatório da interação entre tratamentos e locais, em que: $tl_{ik} \cap N(0, \sigma_{tl}^2)$;

t_i : efeito aleatório referente ao i -ésimo tratamento, tal que $t_i \cap N(0, \sigma_t^2)$;

l_k : efeito aleatório referente ao k -ésimo local, tal que $l_k \cap N(0, \sigma_l^2)$.

Os testes F , para testar os efeitos de tratamentos e de locais, terão o quadrado médio da interação tratamentos e locais como denominador. Não obstante, Banzatto e Kronka (1989) consideram que, se o quadrado médio da interação T×L não for muito diferente do quadrado médio do resíduo, pode-se testar tudo com o quadrado médio do resíduo, o que proporciona maior precisão devido ao maior número de graus de liberdade.

Barbin (2003) comenta que se pode aceitar o modelo inteiramente ao acaso $x_{ijk} = m + t_i + l_k + (tl)_{ik} + e_{ijk}$, para efetuar a análise de um experimento casualizado em blocos, pois a soma dos quadrados do resíduo na análise conjunta é geralmente a soma de quadrados dos resíduos das análises individuais e, muitas vezes, o efeito de blocos dentro de locais, em alguns experimentos, não apresenta importância prática e não precisa aparecer na análise de variância. Entretanto, deve-se respeitar a condição de homoscedasticidade. Porém, no caso dos efeitos de blocos, dentro dos locais, apresentarem importância, o modelo completo é o seguinte:

$$x_{ijk} = m + t_i + b_{k(j)} + l_k + (tl)_{ik} + e_{ijk}$$

$i = 1, \dots, I$ (tratamentos)

$j = 1, \dots, J$ (blocos)

$k = 1, \dots, K$ (locais ou experimentos)

x_{ijk} : valor observado correspondente ao i -ésimo tratamento do j -ésimo bloco do k -ésimo local;

m : média geral;

t_i : efeito referente ao i -ésimo tratamentos;

$b_{k(j)}$: efeito aleatório inerente ao j -ésimo bloco do k -ésimo local, tal que $b_j \cap N(0, \sigma_b^2)$;

l_k : efeito aleatório referente ao k -ésimo local, tal que $l_k \cap N(0, \sigma_l^2)$;

tl_{ik} : efeito da interação entre o i -ésimo tratamento e o k -ésimo local, tal que $tl_{ik} \cap N(0, \sigma_{tl}^2)$;

e_{ijk} : erro experimental associado a x_{ijk} , em que $e_{ijk} \cap N(0, \sigma^2)$ e independentes.

O principal objetivo da análise de grupos de experimentos é efetuar uma análise conjunta e assim obter conclusões válidas para toda uma região, sendo o efeito de local geralmente aleatório. De acordo com Nogueira (1994), eis o quadro de análise de variância (tabela 7.2.2.4), com a casualização em blocos, para o modelo, considerando o efeito de tratamento fixo e o efeito de local aleatório:

Tabela 7.2.2.4. Modelo casualizado em blocos (tratamento: fixo e local: aleatório).

F.V.	G.L.	Q.M.	E(Q.M.)	F
Blocos d. Locais	$K(J-1)$	QM_1	$\sigma^2 + I\sigma_b^2$	
Locais	$(K-1)$	QM_2	$\sigma^2 + J\sigma_{tl}^2 + I\sigma_b^2 + IJ\sigma_l^2$	
Tratamentos (T)	$(I-1)$	QM_3	$\sigma^2 + J\sigma_{tl}^2 + \left(\frac{KJ}{I-1}\right) \sum_{i=1}^I t_i^2$	$\frac{QM_3}{QM_4}$
T×L	$(I-1)(K-1)$	QM_4	$\sigma^2 + J\sigma_{tl}^2$	$\frac{QM_4}{QM_5}$
Resíduo médio	$K(I-1)(J-1)$	QM_5	σ^2	
Total	$IJK-1$			

Em geral, os efeitos de locais são considerados como sendo aleatórios.

b) Hipóteses a serem testadas (teste F) considerando a análise de variância conjunta.

$$b1) H_0 : m_1 = m_2 = \dots = m_I \text{ vs } H_A : m_i \neq m_j \quad i \neq j ;$$

$$b2) H_0 : \sigma_l^2 = 0 \text{ vs } H_A : \sigma_l^2 \neq 0 ;$$

$$b3) H_0 : \sigma_{il}^2 = 0 \text{ vs } H_A : \sigma_{il}^2 \neq 0 .$$

Rejeita-se, ao nível α de significância, $H_0 : m_1 = m_2 = \dots = m_I$, se:

$$\frac{QM_3}{QM_4} = \frac{QM_{TRATAMENTOS(T)}}{QM_{T \times L}} \geq F_{(I-1); (I-1)(K-1); \alpha}$$

Observando a tabela 7.2.2.4, verifica-se a necessidade de uma composição de quadrados médios para encontrar o valor para estudar o efeito de locais. Então:

Rejeita-se, ao nível α de significância, $H_0 : \sigma_l^2 = 0$, se:

$$\frac{QM_2 + QM_5}{QM_1 + QM_4} = \frac{QM_{LOCAIS} + QM_{RESÍDUO MÉDIO}}{QM_{BLOCOS D. LOCAIS} + QM_{T \times L}} \geq F_{n_1; n_2; \alpha}$$

$$n_1 = \frac{(QM_{LOCAIS} + QM_{RESÍDUO})^2}{\frac{(QM_{LOCAIS})^2}{K-1} + \frac{(QM_{RESÍDUO MÉDIO})^2}{K(I-1)(J-1)}} \quad n_2 = \frac{(QM_{BLOCOS D. LOCAIS} + QM_{ESP \times LOCAIS})^2}{\frac{(QM_{BLOCOS D. LOCAIS})^2}{K(J-1)} + \frac{(QM_{T \times L})^2}{(I-1)(J-1)}}$$

Rejeita-se, ao nível α de significância, $H_0 : \sigma_{il}^2 = 0$, se:

$$\frac{QM_4}{QM_5} = \frac{QM_{T \times L}}{QM_{RESÍDUO MÉDIO}} \geq F_{(I-1)(K-1); K(I-1)(J-1); \alpha}$$

7.3. EXEMPLOS ILUSTRATIVOS

Exercício 7.3.1. Considerar a produção de madeira de um ensaio de competição com quatro espécies florestais (E1, E2, E3 e E4), com casualização em cinco blocos, instalados em quatro locais diferentes. A Tabela 7.3.1 apresenta os dados de volume de madeira em m³ por parcela (valores simulados).

Tabela 7.3.1. Volume de madeira das espécies por local.

Local 1					Local 2				
Blocos	E1	E2	E3	E4	Blocos	E1	E2	E3	E4
1	35,7	29,9	42,2	50,4	1	40,4	29,3	42,7	67,2
2	20,6	33,1	48,9	44,9	2	34,9	41,3	23,9	65,3
3	33,0	33,9	38,4	45,0	3	35,0	25,4	36,8	53,0
4	33,4	46,2	28,4	54,6	4	44,6	45,1	33,7	69,4
5	20,9	27,2	37,8	55,1	5	45,1	29,8	47,8	59,4
Total	143,6	170,3	195,7	250,0	Total	200,0	170,9	184,9	314,3

Tabela 7.3.1. Volume de madeira das espécies por local (continuação).

Local 3					Local 4				
Blocos	E1	E2	E3	E4	Blocos	E1	E2	E3	E4
1	40,2	26,1	27,0	51,3	1	28,3	31,2	42,1	48,6
2	42,5	25,5	21,8	41,5	2	23,6	41,8	47,5	48,7
3	49,3	33,7	41,2	55,7	3	27,9	27,6	38,2	51,0
4	39,9	13,1	31,3	64,5	4	34,7	32,8	38,9	38,8
5	49,1	26,0	30,4	56,7	5	13,6	27,9	45,5	49,1
Total	221,0	124,4	151,7	269,7	Total	128,1	161,3	212,2	236,2

a) Análises individuais por local

Tabela 7.3.2. ANOVA para o local 1.

Local 1				
F. V.	G.L.	S.Q.	Q.M.	F
Blocos	4	59,4270	14,8568	0,26 <i>ns</i>
Espécies (E)	3	1234,7000	411,5667	7,16 **
Resíduo	12	689,9450	57,4954	
Total	19	1984,0720		

$$F_{4;12;0,05} = 3,26 \quad F_{3;12;0,01} = 5,95$$

Tabela 7.3.3. ANOVA, no SAS, para o local 1.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	4	59,427000	14,856750	0,26	0,8989
Espécies (E)	3	1234,700000	411,566667	7,16	0,0052
Resíduo	12	689,945000	57,495417		
Total	19	1984,072000			

Tabela 7.3.4. ANOVA para o local 2.

Local 2				
F. V.	G.L.	S.Q.	Q.M.	F
Blocos	4	272,7020	68,1755	1,31 <i>ns</i>
Espécies (E)	3	2.582,1615	860,7205	16,55**
Resíduo	12	624,0860	52,0072	
Total	19	3.478,9495		

$$F_{4;12;0,05} = 3,26 \quad F_{3;12;0,01} = 5,95$$

Tabela 7.3.5. ANOVA, no SAS, para o local 2.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	4	272,702000	68,175500	1,31	0,3208
Espécies (E)	3	2582,161500	860,720500	16,55	0,0001
Resíduo	12	624,086000	52,007167		
Total	19	3478,949500			

Tabela 7.3.6. ANOVA para o local 3.

Local 3				
F. V.	G.L.	S.Q.	Q.M.	F
Blocos	4	341,6730	85,4183	2,27 <i>ns</i>
Espécies (E)	3	2.614,3560	871,4520	23,12**
Resíduo	12	452,3190	37,6933	
Total	19	3.408,3480		

$$F_{4;12;0,05} = 3,26 \quad F_{3;12;0,01} = 5,95$$

Tabela 7.3.7. ANOVA, no SAS, para o local 3.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	4	341,673000	85,418250	2,27	0,1224
Espécies (E)	3	2614,356000	871,452000	23,12	< 0,0001
Resíduo	12	452,319000	37,693250		
Total	19	3408,348000			

Tabela 7.3.8. ANOVA para o local 4.

Local 4				
F. V.	G.L.	S.Q.	Q.M.	F
Blocos	4	87,2930	21,8233	0,59 <i>ns</i>
Espécies (E)	3	1.431,8740	477,2913	12,79**
Resíduo	12	447,6510	37,3043	
Total	19	1.966,8180		

$$F_{4;12;0,05} = 3,26 \quad F_{3;12;0,01} = 5,95$$

Tabela 7.3.9. ANOVA, no SAS, para o local 4.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	4	87,293000	21,823250	0,59	0,6796
Espécies (E)	3	1431,874000	477,291333	12,79	0,0005
Resíduo	12	447,651000	37,304250		
Total	19	1966,818000			

Seja considerado o resumo dos resultados das ANOVAS para os tratamentos (espécies):

Locais	$QM_{RES(k)}$	F
1	57,4954	7,16**
2	52,0072	16,55**
3	37,6933	23,12**
4	37,3043	12,79**

Os resultados apresentados nas tabelas 7.3.1.2, 7.3.1.4, 7.3.1.6, 7.3.1.8 mostram, pelo teste F , que a hipótese $H_0 : m_1 = m_2 = m_3 = m_4$, ao nível de 0,01 de significância, deve ser rejeitada, pois:

$$F = \frac{QM_{ESPÉCIE}}{QM_{RESÍDUO}} > F_{3;12;0,01} = 5,95$$

Para efetuar a análise de variância conjunta para o grupo de experimentos, é necessário testar, considerando os locais, a existência de homoscedasticidade, ou seja, homogeneidade de variâncias.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 \text{ vs } H_A : \sigma_i^2 \neq \sigma_j^2 \quad i \neq j$$

Pelo teste de Hartley:

$$H = \frac{QM_{Max}}{QM_{Min}} = \frac{57,4954}{37,3043} = 1,54$$

Da tabela A9, do apêndice A (Hartley), ocorre que $H_{4;12;0,05} = 4,79$. Logo não há razão para rejeitar $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$. Concluindo-se, pois, pela existência de homogeneidade de variâncias. Desse modo, pode-se efetuar a análise do grupo em estudo.

b) Análise conjunta do grupo de experimentos

Verifica-se que, se a razão entre o maior quadrado de resíduo e o menor for menor do que 4,0, de acordo com Box (1954), os quatro experimentos podem ser reunidos sem restrições em uma única análise conjunta.

Cálculos das somas de quadrados, considerando a análise conjunta:

$$SQ_{TOTAL} = 134.158,05 - \frac{3.134,3^2}{80} = 11.360,0939$$

$$SQ_{BLOCOS \text{ D. LOCAIS}} = \sum_{k=1}^K SQ_{BLOCOS (k)}$$

$$\sum_{k=1}^K SQ_{BLOCOS (k)} = 59,4270 + 272,7020 + 341,6730 + 87,2930 = 761,0950$$

O quadro a seguir apresenta as combinações entre espécies e locais:

	E1	E2	E3	E4	Total
L1	143,6	170,3	195,7	250,0	759,6
L2	200,0	170,9	184,9	314,3	870,1
L3	221,0	124,4	151,7	269,7	766,8
L4	128,1	161,3	212,2	236,2	737,8
Total	692,7	626,9	744,5	1.070,2	3.134,3

$$SQ_{ESPÉCIES} = \frac{(692,7^2 + 626,9^2 + 744,5^2 + 1.070,2^2)}{20} - \frac{3.134,3^2}{80} = 5.824,3034$$

$$SQ_{LOCAIS} = \frac{(759,6^2 + 870,1^2 + 766,8^2 + 737,8^2)}{20} - \frac{3.134,3^2}{80} = 521,9064$$

$$SQ_{ESPÉCIES, LOCAIS} = \frac{(143,6^2 + 170,3^2 + \dots + 212,2^2 + 236,2^2)}{5} - \frac{3.134,3^2}{80}$$

$$SQ_{ESPÉCIES, LOCAIS} = \frac{655.914,8}{5} - \frac{3.134,3^2}{80} = 8.385,0039$$

$SQ_{ESPÉCIES, LOCAIS}$: é a soma de quadrados, considerando as 16 combinações entre espécies e locais.

$$SQ_{ESPÉCIES \times LOCAIS} = SQ_{ESPÉCIES, LOCAIS} - SQ_{ESPÉCIES} - SQ_{LOCAIS}$$

$$SQ_{ESPÉCIES \times LOCAIS} = 8.385,0039 - 5.824,3034 - 521,9064 = 2.038,7881$$

$$SQ_{RESÍDUO \text{ MÉDIO}} = SQ_{TOTAL} - SQ_{LOCAIS} - SQ_{BLOCO \text{ D. LOCAIS}} - SQ_{ESPÉCIES} - SQ_{ESPÉCIES \times LOCAIS}$$

Ou: $SQ_{RESÍDUO \text{ MÉDIO}} = \sum_{k=1}^K SQ_{RESÍDUO (k)}$

$$\sum_{k=1}^K SQ_{RESÍDUO (k)} = 689,9450 + 624,0860 + 452,3190 + 447,6510 = 2.214,0010$$

$$QM_{RESÍDUO \text{ MÉDIO}} = \frac{\sum_{k=1}^K QM_{RESÍDUO (k)}}{K} = \frac{57,4954 + 52,0072 + 37,6932 + 37,3042}{4} = 46,1251$$

De acordo com Barbin (2019), no caso de o efeito de blocos, dentro de locais, não apresentar importância prática e ser desconsiderado na análise de variância, os componentes tratamentos e locais deverão ser testados com a interação Tratamentos×Locais, a qual é testada com o resíduo médio.

No caso de o efeito de blocos dentro de locais – de acordo com Barbin (2003) – não apresentar importância prática, não é necessária a sua inclusão na análise de variância. Dessa maneira, os componentes tratamentos e locais deverão ser testados com a interação Tratamentos×Locais. Contudo, essa interação é testada com o resíduo médio.

Na análise conjunta, quando os ensaios individuais possuem o mesmo número de repetições, o quadrado médio do resíduo é sempre a média aritmética dos resíduos médios das análises individuais.

$$QM_{\text{RESÍDUO MÉDIO}} = \frac{QM_{\text{RES(L1)}} + QM_{\text{RES(L2)}} + \dots + QM_{\text{RES(LK)}}}{K}$$

Quando não forem igualmente repetidos, o quadrado do resíduo médio será uma média ponderada, utilizando os números de graus de liberdade dos ensaios individuais como fator de ponderação.

Para testar o efeito de locais, no caso de o efeito de blocos, dentro de locais, ser considerado importante, tem-se que, considerando os efeitos de locais aleatórios, rejeita-se $H_0 : \sigma_l^2 = 0$, ao nível α de significância, se:

$$F_{\text{LOCAIS}} = \frac{QM_2 + QM_5}{QM_1 + QM_4} = \frac{QM_{\text{LOCAIS}} + QM_{\text{RESÍDUO MÉDIO}}}{QM_{\text{BLOCOS D. LOCAIS}} + QM_{\text{ESPÉCIES} \times \text{LOCAIS}}} \geq F_{n_1; n_2; \alpha}$$

$$F_{\text{LOCAIS}} = \frac{173,9688 + 46,1250}{47,5684 + 226,5320} = \frac{220,0938}{274,1004} = 0,80$$

$$n_1 = \frac{(QM_{\text{LOCAIS}} + QM_{\text{RESÍDUO MÉDIO}})^2}{\frac{(QM_{\text{LOCAIS}})^2}{K-1} + \frac{(QM_{\text{RESÍDUO MÉDIO}})^2}{K(I-1)(J-1)}} = \frac{(173,9688 + 46,1250)^2}{\frac{173,9688^2}{3} + \frac{46,1250^2}{48}} = 5$$

$$n_2 = \frac{(QM_{\text{BLOCOS D. LOCAIS}} + QM_{\text{ESPÉCIES} \times \text{LOCAIS}})^2}{\frac{(QM_{\text{BLOCOS D. LOCAIS}})^2}{K(J-1)} + \frac{(QM_{\text{ESPÉCIES} \times \text{LOCAIS}})^2}{(I-1)(K-1)}}$$

$$n_2 = \frac{(47,5684 + 226,5320)^2}{\frac{47,5684^2}{16} + \frac{226,5320^2}{9}} = 13$$

Não se rejeita $H_0 : \sigma_l^2 = 0$, pois $F_{\text{LOCAIS}} = 0,80 < F_{n_1; n_2; 0,05} = F_{5; 13; 0,05} = 3,03$.

Tabela 7.3.10. ANOVA conjunta para o grupo.

F.V.	G.L.	S.Q.	Q.M.	F
Locais	3	521,9064	---	0,80 <i>ns</i>
Blocos d. Locais	16	761,0950	47,5684	1,03 <i>ns</i>
Espécies (E)	3	5.824,3034	1.941,4345	8,57**
Espécies×Locais	9	2.038,7881	226,5320	4,91 **
Resíduo médio	48	2.214,0010	46,1250	
Total	79	11.360,0940		

$$F_{ESP} = F_{3;9;0,01} = 6,99 \quad F_{LOCAIS} = F_{3;9;0,05} = 3,86 \quad F_{ESP \times LOCAIS} = F_{9;48;0,01} = 2,82$$

Tabela 7.3.11. ANOVA conjunta para o grupo no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Locais	3	521,906375	173,968792	3,77	0,0165
Blocos (Locais)	16	761,095000	47,568438	1,03	0,4430
Espécies	3	5824,303375	1941,434458	42,09	< 0,0001
Locais×Espécies	9	2038,788125	226,532014	4,91	0,0001
Resíduo	48	2214,00100	46,12502		
Total	79	11360,09388			

Testes de hipóteses, usando a interação Local*Espécie como resíduo

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Espécies	3	5824,303375	1941,434458	8,57	0,0053
Locais	3	521,906375	173,968792	0,77	0,5402

O programa SAS, **Prog22**, em anexo no apêndice B, efetua a análise conjunta de grupos de experimentos, entretanto usando os comandos “Test H=Esp E=Local*Esp;” e “Test H=Local E=Local*Esp;” para corrigir os valores do teste *F* referentes aos locais e às espécies. A tabela 7.2.2.4 mostra que os seus quadrados médios devem ser divididos pelo quadrado médio da interação entre os locais e espécies, ou seja, a “interação E×L”.

Se ocorrer a heteroscedasticidade, considerando os locais — o que não é o caso deste exemplo —, pode ser aplicado o método proposto por Cochran (1954), que consiste em ajustar os graus de liberdade do resíduo médio. Logo, a análise de variância conjunta seria realizada, como no caso do quadrado médio do resíduo homogêneo, mas com o número de graus de liberdade ajustado para

consulta das tabelas F e de teste de comparação múltipla, em que n_1^* seria usado para o resíduo médio e n_2^* para os graus de liberdade da interação Locais×Espécies, os quais seriam obtidos da seguinte maneira.

Dado que:

QM_1	QM_2	QM_3	QM_4
57,4954	52,0072	37,6933	37,3043

Então:

$$n_1^* = \frac{[QM_{RES(1)} + QM_{RES(2)} + \dots + QM_{RES(K)}]^2}{\frac{QM_{RES(1)}^2}{glRES(1)} + \frac{QM_{RES(2)}^2}{glRES(2)} + \dots + \frac{QM_{RES(K)}^2}{glRES(K)}}$$

$$n_1^* = \frac{(57,4954 + 52,0072 + 37,6933 + 37,3043)^2}{\frac{57,4954^2}{12} + \frac{52,0072^2}{12} + \frac{37,6933^2}{12} + \frac{37,3043^2}{12}} = 46,30$$

Resultando em $n_1^* = 46$, assim fica atendida a condição:

$$glRES(K) < n_1^* < \sum_{k=1}^K glRES(k), \text{ pois } (12 < n_1^* = 46 < 48).$$

$glRES(k)$: número de graus de liberdade do resíduo inerente ao k -ésimo experimento.

Sendo $n_1^* = 46$, segundo Cochran (1954), o número de graus de liberdade do resíduo médio, para a consulta da tabela F e testes de comparações múltiplas, é próximo ao que foi obtido na análise de variância, considerando homogeneidade de variâncias, que é igual a 48. A seguir o cálculo de n_2^* para os graus de liberdade da interação Locais×Espécies:

$$V_1 = \frac{\sum_{k=1}^K QM_{RES(k)}}{K} = \frac{184,5002}{4} = 46,1251$$

$$V_2 = \frac{\sum_{k=1}^K QM_{RES(k)}^2}{K} = \frac{57,4954^2 + 52,0072^2 + 37,6933^2 + 37,3043^2}{4} = \frac{8.822,8655}{4}$$

$$V_2 = \frac{\sum_{k=1}^K QM_{RES(k)}^2}{K} = 2.205,7164$$

$$n_2^* = \frac{(I-1)(K-1)^2 V_1^2}{(K-2)V_2 + V_1^2} = \frac{(4-1)(4-1)^2 \times 46,1251^2}{(4-2) \times 2.205,7164 + 46,1251^2} = 8,8$$

I : é o número de tratamentos;

K : é o número de locais;

$QM_{Res(k)}$: quadrado médio residual correspondente ao k -ésimo local ou experimento.

Então $n_2^* = 9$ é o número de graus de liberdade da interação de espécies×locais para consulta da tabela F e testes de comparações múltiplas. Esse valor é o mesmo obtido na montagem do quadro de análise de variância, tabela 7.3.10 do exercício 7.3.1, na análise estruturada, considerando a homogeneidade de variâncias.

Tendo em vista que a interação Espécies×Locais é significativa, isto é, o comportamento das espécies difere entre os locais. Por outro lado, como o valor de F também foi significativo para o tratamento espécie, existe pelo menos uma espécie que teve um comportamento melhor nos locais, justificando realizar um teste de comparação múltipla de médias dos tratamentos.

Estudo de cada espécie dentro dos locais, o que corresponde em desdobrar os graus de liberdade de locais mais o da interação Espécies×Locais.

	E1	E2	E3	E4	Total
L1	143,6	170,3	195,7	250,0	759,6
L2	200,0	170,9	184,9	314,3	870,1
L3	221,0	124,4	151,7	269,7	766,8
L4	128,1	161,3	212,2	236,2	737,8
Total	692,7	626,9	744,5	1.070,2	3.134,3

$$SQ_{E1 D. LOCAIS} = \frac{1}{5}(143,6^2 + 200,0^2 + 221,0^2 + 128,1^2) - \frac{692,7^2}{20} = 1.182,6495$$

$$SQ_{E2 D. LOCAIS} = \frac{1}{5}(170,3^2 + 170,9^2 + 124,4^2 + 161,3^2) - \frac{626,9^2}{20} = 290,2095$$

$$SQ_{E3 D. LOCAIS} = \frac{1}{5}(195,7^2 + 184,9^2 + 151,7^2 + 212,2^2) - \frac{744,5^2}{20} = 391,6335$$

$$SQ_{E4 D. LOCAIS} = \frac{1}{5}(250,0^2 + 314,3^2 + 269,7^2 + 236,2^2) - \frac{1.070,2^2}{20} = 696,2020$$

Tabela 7.3.12. ANOVA de cada espécie dentro dos locais.

Contrastes	G.L.	S.Q.	Q.M.	F
Locais	3	521,9064	-	-
Espécie 1 dentro de Locais	3	1.182,6495	394,2165	8,55**
Espécie 2 dentro de Locais	3	290,2095	96,7365	2,10 <i>ns</i>
Espécie 3 dentro de Locais	3	391,6335	130,5445	2,83*
Espécie 4 dentro de Locais	3	696,2020	232,0673	5,03**
Resíduo médio	48	2.214,0010	46,1250	

$$F_{3; 48; 0,05} = 2,81 \quad F_{3; 48; 0,01} = 4,24$$

Tabela 7.3.13. ANOVA de cada espécie dentro dos locais no SAS.

Desdobramento de cada espécie dentro dos locais					
Espécies	G.L.	S.Q.	Q.M.	F	Pr > F
E1	3	1182,649500	394,216500	8,55	0,0001
E2	3	290,209500	96,736500	2,10	0,1129
E3	3	391,633500	130,544500	2,83	0,0482
E4	3	696,202000	232,067333	5,03	0,0041

Observar que:

$$SQ_{E1 D. LOCAIS} + SQ_{E2 D. LOCAIS} + SQ_{E3 D. LOCAIS} + SQ_{E4 D. LOCAIS} = SQ_{LOCAIS} + SQ_{ESP \times LOCAIS}$$

$$SQ_{E1 D. LOCAIS} + SQ_{E2 D. LOCAIS} + SQ_{E3 D. LOCAIS} + SQ_{E4 D. LOCAIS} = 1.182,6495 + \dots + 696,2020$$

$$SQ_{E1 D. LOCAIS} + SQ_{E2 D. LOCAIS} + SQ_{E3 D. LOCAIS} + SQ_{E4 D. LOCAIS} = 2.560,6945$$

$$SQ_{LOCAIS} + SQ_{ESP \times LOCAIS} = 521,9064 + 2.038,7881 = 2.560,6945$$

Os resultados (tabela 7.3.12) mostram que o comportamento da espécie E2, para o nível de significância $\alpha = 0,05$, foi não significativo nos quatro locais. Entretanto, como as demais espécies apresentaram comportamentos significantes, pode-se efetuar o teste de Tukey para as espécies E1, E3 e E4.

Dado o quadro auxiliar:

	E1	E2	E3	E4	Total
L1	143,6	170,3	195,7	250,0	759,6
L2	200,0	170,9	184,9	314,3	870,1
L3	221,0	124,4	151,7	269,7	766,8
L4	128,1	161,3	212,2	236,2	737,8
Total	692,7	626,9	744,5	1.070,2	3.134,3

	Médias		
Locais	E1	E3	E4
Local 1	28,72	39,14	50,00
Local 2	40,00	36,98	62,86
Local 3	44,20	30,34	53,94
Local 4	25,62	42,44	47,24

$$\Delta = q_{4; 48; 0,05} \times \frac{\sqrt{QM_{RESÍDUO MÉDIO}}}{\sqrt{5}} = 3,77 \times \frac{\sqrt{46,1250}}{\sqrt{5}} = 11,45$$

Tabela 7.3.14. Teste de Tukey das espécies E1, E3 e E4 nos locais.

	Médias				
	E1		E3		E4
Local 3	44,20 a	Local 4	42,44 a	Local 2	62,86 a
Local 2	40,00 ab	Local 1	39,14 ab	Local 3	53,94 ab
Local 1	28,72 bc	Local 2	36,98 ab	Local 1	50,00 b
Local 4	25,62 c	Local 3	30,34 b	Local 4	47,24 b

A espécie E1 teve o melhor desempenho nos locais 2 e 3. A espécie E3 apresentou o melhor desempenho nos locais 1, 2 e 4. A espécie E4 mostrou o melhor desempenho nos locais 2 e 3.

Exercício 7.3.2. Conforme o exemplo apresentado em Barbin (2003), os dados se referem à altura média (em m) de 25 plantas de *Eucalyptus grandis*, com sete anos de idade (em 1980) em três locais (L), com casualização em quatro blocos de seis progênies (tratamentos). Nos quadros abaixo, estão apresentados os dados para: os locais L1 (Araraquara), L2 (Bento Quirino) e L3 (Mogi Guaçu), considerando as seis progênies (P1, P2,..., P6) para os quatro blocos. Fonte: Instituto Florestal (Tupi-SP).

Tabela 7.3.15. Dados das seis progênies para os três locais.

Ensaio em Araraquara					
Trat.	Bloco1	Bloco2	Bloco3	Bloco4	Total
P1	20,3	19,6	23,5	19,1	82,5
P2	21,7	19,3	16,7	18,5	76,2
P3	22,0	24,9	24,4	20,8	92,1
P4	20,8	23,0	21,3	24,9	90,0
P5	21,5	22,3	22,1	21,9	87,8
P6	19,6	17,7	18,7	22,0	78,0
Totais	125,9	126,8	126,7	127,2	506,6

Ensaio em Bento Quirino					
Trat.	Bloco1	Bloco2	Bloco3	Bloco4	Total
P1	10,2	11,7	9,1	8,1	39,1
P2	16,1	10,8	10,9	10,3	48,1
P3	17,7	13,1	14,2	11,0	56,0
P4	13,5	14,4	11,2	12,8	51,9
P5	20,5	12,5	11,3	12,2	56,5
P6	12,0	13,0	12,3	10,6	47,9
Totais	90,0	75,5	69,0	65,0	299,5

Ensaio em Mogi Guaçu					
Trat.	Bloco1	Bloco2	Bloco3	Bloco4	Total
P1	22,7	21,4	22,9	22,0	89,0
P2	22,6	21,4	20,7	20,8	85,5
P3	21,4	21,7	22,5	19,4	85,0
P4	25,0	23,6	23,3	24,8	96,7
P5	26,4	26,4	28,0	27,3	108,1
P6	20,6	23,5	19,4	21,9	85,4
Totais	138,7	138,0	136,8	136,2	549,7

Tabela 7.3.16. Análises (quadrados médios) por local.

		Q.M.		
F.V.	G.L.	Araraquara	Bento Quirino	Mogi Guaçu
Blocos	3	0,0494	20,0660	0,2138
Progênes	5	10,7007	10,4324	21,2114
Resíduo	15	3,7831	3,7560	1,4354
Total	23			

$$F_{Max} = \frac{QM_{Max}}{QM_{Min}} = \frac{3,7831}{1,4354} = 2,63$$

Considerando que a relação entre o maior e menor quadrado do resíduo é inferior a 4, os três ensaios poderão ser reunidos numa única análise conjunta (Box, 1954).

Nessa análise, Barbin (2003) considerou o modelo inteiramente ao acaso, $x_{ijk} = m + t_i + l_k + (tl)_{ik} + e_{ijk}$, para efetuar a análise desse experimento casualizado em blocos. O autor considerou que o efeito de blocos dentro de locais não apresentou importância prática, portanto não deve aparecer na análise de variância. Porém, foi respeitada a condição de homoscedasticidade, resultando na seguinte análise de variância conjunta para os três experimentos.

Os testes F para tratamentos (T) e locais (L) são feitos com o quadrado médio da interação $T \times L$ no denominador. A interação $T \times L$ é testada com o resíduo médio. Na tabela 7.3.17, está a análise de variância, apresentada por Barbin (2003), na qual foi retirado o efeito de blocos dentro de locais.

Tabela 7.3.17. Análise conjunta para os três experimentos.

F.V.	G.L.	S.Q.	Q.M.	F
Progênes (T)	5	135,1511	27,0302	3,53 *
Locais (L)	2	1.490,9453	745,4726	97,36 **
Interação $T \times L$	10	76,5714	7,6571	2,56 *
Resíduo Médio	45		2,9915	

$$\text{Resíduo Médio} = \frac{3,7831 + 3,7560 + 1,4354}{3} = 2,9915$$

$$F_{5; 10; 0,05} = 3,33 \quad F_{2; 10; 0,01} = 7,56$$

$$F_{10; 45; 0,05} = 2,06$$

Considerando os resultados da tabela 7.3.17, Barbin (2003) menciona que, como a interação $T \times L$ é significativa, isto quer dizer que o comportamento dos tratamentos difere entre os locais. Então, o ideal é desdobrar os graus de liberdade e estudar os tratamentos (progênes) dentro de cada local, tendo a vantagem sobre as análises individuais de contar com o maior número de graus de liberdade para o resíduo.

Considerar a análise de variância de cada espécie dentro dos locais, em que L1 (Araraquara), L2 (Bento Quirino) e L3 (Mogi Guaçu):

Tabela 7.3.18. ANOVA de cada progênie dentro dos locais.

F.V.	G.L.	S.Q.	Q.M.	F
Locais (L)	2	1.490,9453		
Progênies d. L1	5	53,5033	10,7007	3,58 **
Progênies d. L2	5	52,1621	10,4324	3,49 **
Progênies d. L3	5	106,0571	21,2114	7,09 **
Resíduo Médio	45	----	2,9915	

$$F_{5; 45; 0,05} = 2,39 \quad F_{5; 45; 0,01} = 3,38$$

Tabela 7.3.19. ANOVA de cada espécie dentro dos locais no SAS.

Desdobramento de cada espécie dentro dos locais					
Local	G.L	S.Q.	Q.M.	F	Pr > F
L1	5	53,503333	10,700667	3,58	0,0083
L2	5	52,162083	10,432417	3,49	0,0095
L3	5	106,057083	21,211417	7,09	< 0,0001

Aplicando o teste de Tukey:

$$\Delta = q_{5; 45; 0,05} \sqrt{\frac{QM_{\text{RESÍDUO MÉDIO}}}{4}} = 4,21 \times \sqrt{\frac{2,9915}{4}} = 3,64$$

Tabela 7.3.20. Teste de Tukey de cada progênie dentro dos locais.

Dentro de Araraquara		Dentro de Bento Quirino		Dentro de Mogi Guaçu	
$\hat{m}_4$	22,50 a	$\hat{m}_5$	14,12 a	$\hat{m}_5$	27,02 a
$\hat{m}_5$	21,95 a	$\hat{m}_3$	14,00 a	$\hat{m}_4$	24,18 ab
$\hat{m}_1$	20,62 a	$\hat{m}_4$	12,98 ab	$\hat{m}_1$	22,25 b
$\hat{m}_6$	19,50 a	$\hat{m}_2$	12,02 ab	$\hat{m}_2$	21,38 b
$\hat{m}_2$	19,05 a	$\hat{m}_6$	11,98 ab	$\hat{m}_6$	21,35 b
$\hat{m}_3$	14,12 b	$\hat{m}_1$	9,78 b	$\hat{m}_3$	21,25 b

Médias seguidas pela mesma letra, ao nível de significância $\alpha = 0,05$, não diferem pelo teste de Tukey.

Observar que a progênie $P3 (\hat{m}_3)$ teve o pior comportamento em Araraquara, enquanto em Bento Quirino apresentou um comportamento muito melhor. O comportamento da progênie $P1 (\hat{m}_1)$ também foi melhor em Araraquara, porém pior em Bento Quirino, assim explicando algumas causas da interação $T \times L$ ter dado significativo.

Por outro lado, efetuando a análise dos dados apresentados por Barbin (2003), mas considerando o efeito de blocos dentro dos locais, há, nas tabelas 7.3.21, 7.3.22 e 7.3.23, a análise conjunta dos três experimentos.

De acordo com os resultados apresentados na tabela 7.3.21, rejeita-se, ao nível de significância de $\alpha = 0,05$, a hipótese de igualdade de médias, $H_0 : m_1 = m_2 = \dots = m_6$, para as seis progênes.

Tabela 7.3.21. ANOVA considerando o efeito de blocos d. Locais.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos d. Locais	9	60,9875	6,7764	2,27 *
Progênes (T)	5	135,1511	27,0302	3,53 *
Locais (L)	2	1.490,9453	745,4726	(97,36)**
Interação $T \times L$	10	76,5714	7,6571	2,56 *
Resíduo Médio	45	134,6175	2,9915	
Total	71	1.898,2728		

Tabela 7.3.22. ANOVA, no SAS, considerando o efeito de blocos d. Locais.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos (Locais)	9	60,987500	6,776389	2,27	0,0345
Locais	2	1490,945278	745,472639	249,20	< 0,0001
Progênes	5	135,151111	27,030222	9,04	< 0,0001
Locais×Progênes	10	76,571389	7,657139	2,56	0,0153
Resíduo	45	134,617500	2,991500		
Total	71	1898,272778			

Tabela 7.3.23. ANOVA com as correções dos valores de F para progênes e locais.

Testes de Hipóteses, usando a interação Locais*Progênes como resíduo					
F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Progênes	5	135,151111	27,030222	3,53	0,0425
Locais	2	1490,945278	745,472639	97,36	< 0,0001

O programa SAS, **Prog23**, em anexo no apêndice B, efetua a análise conjunta de grupos de experimentos, em que, usando os comandos “Test H=PROG E=Local*PROG;” e “Test H=Local E=Local*PROG;”, os valores do teste F referentes aos locais e às progêneses são corrigidos. A tabela 7.3.2.3 apresenta os seus quadrados médios que devem ser divididos pelo quadrado médio da interação entre os locais (L) e as progêneses (T), isto é, “interação $T \times L$ ”.

Por outro lado, considerando o efeito de blocos dentro de locais importante, para testar o efeito de locais, ao nível α de significância:

$$F_{LOCAIS} = \frac{QM_{LOCAIS} + QM_{RESÍDUO MÊDIO}}{QM_{BLOCOS D. LOCAIS} + QM_{PROGÊNES \times LOCAIS}} \geq F_{n_1; n_2; \alpha}$$

$$F_{LOCAIS} = \frac{745,4726 + 2,9915}{6,7764 + 7,6571} = 51,86$$

$$n_1 = \frac{(QM_{LOCAIS} + QM_{RESÍDUO MÊDIO})^2}{\frac{(QM_{LOCAIS})^2}{K-1} + \frac{(QM_{RESÍDUO MÊDIO})^2}{K(I-1)(J-1)}} = \frac{(745,4726 + 2,9915)^2}{\frac{745,4726^2}{2} + \frac{2,9915^2}{45}} = 2$$

$$n_2 = \frac{(QM_{BLOCOS D. LOCAIS} + QM_{PROGÊNES \times LOCAIS})^2}{\frac{(QM_{BLOCOS D. LOCAIS})^2}{K(J-1)} + \frac{(QM_{PROGÊNES \times LOCAIS})^2}{(I-1)(K-1)}}$$

$$n_2 = \frac{(6,7764 + 7,6571)^2}{\frac{6,7764^2}{9} + \frac{7,6571^2}{10}} = 19$$

$$F_{n_1; n_2; 0,05} = F_{2; 19; 0,05} = 3,52$$

Dado que o valor calculado da estatística F para local foi 51,86, há razão para rejeitar a hipótese de nulidade para o efeito de locais, ao nível de significância $\alpha = 0,05$, isto é, existe influência dos locais nos crescimentos das progêneses. Entretanto, como a interação $T \times L$ é significativa, isso quer dizer que o comportamento dos tratamentos difere de acordo com os locais, então o ideal é desdobrar, como já realizado, os graus de liberdade, estudando os tratamentos dentro de cada local.

Piedade (1987) apresenta os procedimentos para realizar a análise de grupos de experimentos em faixas. Os experimentos em faixas estão apresentados no item 8.2 do capítulo 8.

7.4. BIBLIOGRAFIA

BANZATTO, D. A.; KRONKA, N. **Experimentação agrícola**. Jaboticabal: FUNEP, 1989. 247 p.

BARBIN, D. **Planejamento e análise estatística de experimentos agronômicos**. Arapongas/PR: Editora Midas, 2003. 208 p.

BARBIN, D. **Componentes de variância. Teoria e aplicações**. 3^a ed. Piracicaba/SP: FEALQ, 2019. 144 p.

BOX, G.E.P. Some theorems on quadratic forms applied in the study of analysis variance problems, I. **Ann. Math. Stat.** v.25, p. 290-302, 1954.

COCHRAN, W. G. The combination of estimates from different experiments. **Biometrics**, v.10, p.101-129, 1954.

NOGUEIRA, M. C. S. **Curso de estatística experimental aplicada à experimentação agronômica**. Piracicaba/SP: ESALQ/USP, 1994. 247 p.



CAPÍTULO 8

ENSAIOS COM PARCELAS SUBDIVIDIDAS (*SPLIT-PLOT*)

Este capítulo apresenta definições, conceitos e aplicações sobre os arranjos com parcelas subdivididas. Trata sobre o ensaio com parcelas subdivididas na sua conceituação original e experimentos em faixas. Exibe as análises univariada e multivariada nos experimentos com observações no tempo. Apresenta, ainda, os componentes de variância, desdobramento dos graus de liberdade quando a interação dos fatores A e B é significativa e o uso do teste de Tukey na comparação dos tratamentos.

8.1 EXPERIMENTOS EM PARCELAS SUBDIVIDIDAS

Todas as combinações entre os níveis dos fatores que denotam os tratamentos nos ensaios fatoriais, sem qualquer restrição e independentemente do delineamento escolhido, são casualizadas nas parcelas. Os ensaios em parcelas subdivididas são utilizados quando se deseja estudar simultaneamente dois grupos de tratamentos, porém, com uma estruturação um pouco diferente dos ensaios fatoriais. A principal característica das parcelas subdivididas, devido à natureza dos fatores, são as dificuldades ligadas à instalação do experimento no campo e na casualização dos dois grupos ou fatores.

Nos ensaios fatoriais, a casualização de todas as combinações é feita de acordo com o princípio do delineamento utilizado, isto é, inteiramente ao acaso, casualizado em blocos e quadrado latino. Na parcela subdividida, a casualização é feita em duas etapas. Primeiramente casualizam-se os níveis do fator *A*, que será avaliado nas parcelas, seguindo as diretrizes do delineamento em uso, sendo os delineamentos inteiramente ao acaso e os casualizados em blocos os mais utilizados. Em uma segunda etapa, são sorteados dentro de cada parcela os níveis do fator *B* que será avaliado nas subparcelas, isto é, cada parcela atua como se fosse um bloco para os tratamentos das subparcelas.

Silva e Silva (1982) descrevem um experimento casualizado em “*J*” blocos, arranjado em parcelas subdivididas, em que o interesse é testar três espécies florestais (E1, E2 e E3) de *Pinus sp*, em três espaçamentos distintos (I, II e III), tal que o objetivo principal é obter mais informações sobre o comportamento das espécies. O delineamento apresenta a seguinte distribuição: os espaçamentos ficam dispostos nas parcelas (blocos) e as espécies casualizadas dentro de cada espaçamento:

Tratamento A: espaçamentos = I, II e III (distribuídos aleatoriamente dentro de cada bloco) e Tratamento B: espécies = E1, E2 e E3 (distribuídas aleatoriamente dentro de cada espaçamento). A exemplo, para a distribuição das combinações, o seguinte quadro:

Bloco1	IIIE2	IIIE1	IIIE3	IE1	IE3	IE2	IIIE1	IIIE2	IIIE3
Bloco2	IE3	IE2	IE1	IIIE3	IIIE1	IIIE2	IIIE2	IIIE1	IIIE3
...	...	...	...	...	...	...	...	...	...
Bloco J	IIIE1	IIIE2	IIIE3	IIIE3	IIIE2	IIIE1	IE2	IE3	IE1

De acordo com Cochran e Cox (1957), a opção pela aplicação do experimento em parcelas subdivididas é mais vantajosa quando os efeitos dos tratamentos do fator *B* (subparcelas) e da interação dos fatores *A* e *B* apresentam maior interesse que os efeitos do fator *A*, pois o aumento da precisão nas estimativas dos efeitos do fator *B* e da interação entre *A* e *B* são obtidas mediante a redução na precisão das estimativas dos tratamentos do fator *A*.

Os níveis do fator *A* distribuídos nas parcelas são denominados de tratamentos principais e os níveis do fator *B* colocados nas subparcelas são chamados de tratamentos secundários. Esse esquema resulta que os tratamentos primários são confundidos com as parcelas, enquanto os tratamentos secundários não são confundidos. Assim sendo, deve-se designar o fator de maior interesse, quando factível, como os tratamentos secundários. Há, então, dois resíduos distintos:

Resíduo (a): variabilidade residual estudada nas parcelas (tratamentos principais) e utilizada na comparação dos níveis do fator A ;

Resíduo (b): variabilidade residual entre as subparcelas dentro das parcelas e utilizado na comparação dos níveis do fator B (tratamentos secundários).

Em alguns casos, na pesquisa em ciências agrárias, esses ensaios são vantajosos no que se refere à praticidade de instalação. É o caso, por exemplo, de quando se deseja avaliar o efeito de diferentes intensidades de irrigação, diferentes intensidades de queimadas, vários tipos de podas etc.

Em algumas situações especiais, as subparcelas podem também ser subdivididas em subsubparcelas, originando três diferentes resíduos: Resíduo (a), inerente às parcelas; Resíduo (b), correspondente às subparcelas; e o Resíduo (c), associado às subsubparcelas. Desse modo, o processo de subdivisão pode evoluir para outras subdivisões.

No caso de parcelas subdivididas no espaço, verifica-se a subdivisão da área (parcela) em subáreas (subparcelas). Porém, quando há medição de tratamentos no tempo, cada ocasião medida é considerada uma subparcela.

Os experimentos em parcelas subdivididas são menos eficientes em termos de precisão estatística que os ensaios fatoriais, pois se verifica a ocorrência de dois resíduos. Consequentemente, há a partição do número de graus de liberdade do resíduo em dois componentes, denominados de graus de liberdade do resíduo (a) e graus de liberdade do resíduo (b), acarretando o aumento do erro experimental. Desse modo, verifica-se que os efeitos são avaliados, dentro dos experimentos com parcelas subdivididas, com menor precisão que nos ensaios fatoriais correspondentes.

Por outro lado, sempre que possível, deve-se optar pelos ensaios fatoriais e somente utilizar os delineamentos em parcelas subdivididas quando resultar em ganho de natureza operacional ou em economia de recursos materiais e financeiros. Deve-se sempre buscar alocar o fator de tratamento de maior importância, ou que requer menor quantidade de material nas subparcelas, tendo em vista que o erro experimental das subparcelas tende a ser menor, haja vista que o número de graus de liberdade do resíduo (b) é maior que o número de graus de liberdade do resíduo (a).

Quando o efeito da interação entre os tratamentos primários e os tratamentos secundários for significativo por meio do teste F , isso indica que os tratamentos primários possuem um comportamento diferenciado na presença de cada um dos tratamentos secundários, e/ou os tratamentos secundários apresentam um comportamento diferenciado na presença de cada um dos tratamentos primários. Nesses casos, o procedimento a ser adotado é a decomposição dos graus de liberdade da interação $A \times B$ mais os graus de liberdade dos tratamentos primários ou dos tratamentos secundários.

a) Quando a decomposição dos graus de liberdade da interação $A \times B$ ocorrer com a soma dos graus de liberdade dos tratamentos secundários, estuda-se o efeito dos tratamentos secundários na presença de cada tratamento primário, em que a é o número de níveis do fator A . Assim:

$$SQ_{Bd..A_1} + SQ_{Bd..A_2} + \dots + SQ_{Bd..A_a} = SQ_B + SQ_{A \times B}$$

Nesse caso, o denominador apropriado para compor o teste F , para testar as hipóteses, será o próprio $QM_{RES(b)}$.

b) Quando a decomposição dos graus de liberdade da interação $A \times B$ ocorrer com a soma dos graus de liberdade dos tratamentos primários, estuda-se o efeito dos tratamentos primários na presença de cada tratamento secundário, em que b é o número de níveis do fator B . Assim:

$$SQ_{Ad.B_1} + SQ_{Ad.B_2} + \dots + SQ_{Ad.B_b} = SQ_A + SQ_{A \times B}$$

No caso da análise dos efeitos dos tratamentos primários na presença de cada tratamento secundário, o quadrado médio do resíduo a ser usado será o obtido pela expressão:

$$QM_{RES(MÉDIO)} = \frac{QM_{RES(a)} + (b-1)QM_{RES(b)}}{b}$$

Em que o $QM_{RES(MÉDIO)}$ está associado a (n'_{gl}) graus de liberdade, tal que n'_{gl} é obtido pela fórmula de Satterthwaite (1946):

$$n'_{gl} = \frac{[QM_{RES(a)} + (b-1)QM_{RES(b)}]^2}{\frac{[QM_{RES(a)}]^2}{g.l. Res(a)} + \frac{(b-1)^2 [QM_{RES(b)}]^2}{g.l. Res(b)}}$$

Tal que $g.l. Res(a)$ é o número de graus de liberdade associado ao $QM_{RES(a)}$ e $g.l. Res(b)$ se refere aos graus de liberdade de $QM_{RES(b)}$, em que J é o número de repetições.

Considerando o delineamento com casualização em blocos e os fatores A e B , seja o modelo matemático:

$$x_{ikj} = m + \theta_j + \alpha_i + (\alpha\theta)_{ij} + \beta_k + (\alpha\beta)_{ik} + e_{ikj}$$

x_{ikj} : valor da variável resposta da combinação correspondente ao nível (i) do fator A e ao nível (k) do fator B no bloco (j) , tal que: $i = 1, \dots, a$; $k = 1, \dots, b$ e $j = 1, \dots, J$, em que a é o número de níveis do fator A e b é o número de níveis do fator B .

m : média geral;

θ_j : considerando o efeito aleatório do j -ésimo bloco, tal que $N(0, \sigma_\theta^2)$;

θ_j : efeito fixo do j -ésimo bloco, tal que $E(\theta_j) = \theta_j$, $E(\theta_j^2) = \theta_j^2$ e $\sum_{j=1}^J \theta_j = 0$;

α_i : efeito do i -ésimo nível do fator A (parcela);

$(\alpha\theta)_{ij}$: erro da parcela (ij) , tal que $N(0, \sigma_{\alpha\theta}^2)$;

β_k : efeito do k -ésimo nível do fator B (subparcela);

$(\alpha\beta)_{ik}$: efeito da interação $A_i \times B_k$;

e_{ikj} : erro da subparcela (ikj) , tal que $N(0, \sigma^2)$.

Nas tabelas 8.1.1 e 8.1.2, de acordo com Steel e Torrie (1960), considerando o experimento em parcelas subdivididas casualizado em blocos, em que o efeito de blocos é aleatório, estão apresentados os componentes de variância para os modelos fixo, aleatório e mistos.

Tabela 8.1.1. Componentes de variância, considerando casualizado em blocos.

F.V.	G.L.	E(Q.M.)	
		<i>A</i> e <i>B</i> fixos:	<i>A</i> e <i>B</i> aleatórios:
Blocos	$J - 1$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + ab\sigma_\theta^2$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + ab\sigma_\theta^2$
Trat. (A)	$a - 1$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + Jb \frac{\sum_{i=1}^a \alpha_i^2}{a-1}$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + J\sigma_{\alpha\beta}^2 + Jb\sigma_\alpha^2$
Resíduo (a)	$(a-1)(J-1)$	$\sigma^2 + b\sigma_{\alpha\theta}^2$	$\sigma^2 + b\sigma_{\alpha\theta}^2$
Parcelas	$(aJ - 1)$		
Trat. (B)	$b - 1$	$\sigma^2 + Ja \frac{\sum_{k=1}^b \beta_k^2}{b-1}$	$\sigma^2 + J\sigma_{\alpha\beta}^2 + Ja\sigma_\beta^2$
Interação A×B	$(a-1)(b-1)$	$\sigma^2 + J \frac{\sum_{i,k} (\alpha\beta)_{ik}^2}{(a-1)(b-1)}$	$\sigma^2 + J\sigma_{\alpha\beta}^2$
Resíduo (b)	$a(b-1)(J-1)$	σ^2	σ^2
Total	$(abJ - 1)$		

Tabela 8.1.2. Componentes de variância com casualização em blocos (modelos mistos).

F.V.	E(Q.M.)	
	<i>A</i> (aleatório) e <i>B</i> fixo	<i>A</i> (fixo) e <i>B</i> (aleatórios)
Blocos	$\sigma^2 + b\sigma_{\alpha\theta}^2 + ab\sigma_\theta^2$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + ab\sigma_\theta^2$
Trat. (A)	$\sigma^2 + b\sigma_{\alpha\theta}^2 + Jb\sigma_\alpha^2$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + J \frac{a}{a-1} \sigma_{\alpha\beta}^2 + Jb \frac{\sum_{i=1}^a \alpha_i^2}{a-1}$
Resíduo (a)	$\sigma^2 + b\sigma_{\alpha\theta}^2$	$\sigma^2 + b\sigma_{\alpha\theta}^2$
Parcelas		
Trat. (B)	$\sigma^2 + J \frac{b}{b-1} \sigma_{\alpha\beta}^2 + Ja \frac{\sum_{k=1}^b \beta_k^2}{b-1}$	$\sigma^2 + Ja\sigma_\beta^2$
Interação A×B	$\sigma^2 + J \frac{b}{b-1} \sigma_{\alpha\beta}^2$	$\sigma^2 + J \frac{a}{a-1} \sigma_{\alpha\beta}^2$
Resíduo (b)	σ^2	σ^2
Total		

Exercício 8.1.1. Considerar o experimento em que há um plantio com uma espécie florestal, no qual serão comparados quatro métodos de manejo $M1$, $M2$, $M3$ e $M4$ (Fator M), aplicados em parcelas de tamanho de um hectare e distribuídos em quatro blocos. Em cada parcela de um hectare foram alocados, considerando subparcelas de 0,25 hectares, quatro diferentes tipos de adubação (Fator A), o que caracteriza o uso do arranjo de parcelas subdivididas. O quadro abaixo mostra o croqui com a distribuição dos fatores M (métodos de manejo) e A (tipos de adubação). A tabela 8.1.3 apresenta os valores dos volumes (dados simulados) das combinações dos fatores M e A .

Bloco1				Bloco2			
M1A2	M3A4	M2A3	M4A2	M4A2	M2A3	M3A4	M1A1
M1A4	M3A1	M2A2	M4A1	M4A4	M2A4	M3A1	M1A2
M1A1	M3A3	M2A4	M4A1	M4A1	M2A1	M3A3	M1A4
M1A3	M3A2	M2A1	M4A4	M4A3	M2A2	M3A2	M1A3
Bloco3				Bloco4			
M1A3	MAB2	M2A3	M3A1	M4A3	M1A3	M3A2	M2A1
M1A2	M4A3	M2A1	M3A4	M4A4	M1A1	M3A1	M2A3
M1A4	M4A4	M2A4	M3A2	M4A2	M1A2	M3A4	M2A4
M1A1	M4A1	M2A2	M3A3	M4A1	M1A4	M3A3	M2A2

Tabela 8.1.3. Volumes em m^3 por parcela das combinações dos fatores M e A .

Fator M	Fator A	Blocos				Totais
		1	2	3	4	
M1	A1	58,1	58,4	39,5	44,6	200,6
M1	A2	72,7	77,9	58,5	61,7	270,8
M1	A3	65,9	71,7	54,2	55,5	247,3
M1	A4	59,9	55,7	37,7	46,2	199,5
M2	A1	71,5	92,7	60,5	46,7	271,4
M2	A2	76,7	93,7	56,5	69,1	296,0
M2	A3	79,7	87,7	55,1	60,5	283,0
M2	A4	85,4	76,5	58,7	68,7	289,3
M3	A1	83,0	77,9	59,4	67,0	287,3
M3	A2	84,5	67,1	59,9	62,2	273,7
M3	A3	85,9	64,4	83,4	67,0	300,7
M3	A4	85,7	74,7	72,2	69,0	301,6
M4	A1	99,5	87,4	71,9	70,2	329,0
M4	A2	93,7	89,7	76,7	77,9	338,0
M4	A3	91,0	87,0	63,7	67,9	309,6
M4	A4	95,4	92,0	75,4	63,1	325,9
Totais		1288,6	1254,5	983,3	997,3	4523,7

$$C = \frac{G^2}{IJ} = \frac{4.523,7^2}{64} = 319.747,8389$$

$$SQ_{TOTAL} = 58,1^2 + 72,7^2 + \dots + 67,9^2 + 63,1^2 - C$$

$$SQ_{TOTAL} = 333.000,19 - 319.747,84 = 13.252,3511$$

$$SQ_{BLOCOS} = \frac{1}{16}(1.288,6^2 + 1.254,5^2 + 983,3^2 + 997,3^2) - C = 4.986,3105$$

	Bloco1	Bloco2	Bloco3	Bloco4	Totais
M1	256,6	263,7	189,9	208,0	918,2
M2	313,3	350,6	230,8	245,0	1.139,7
M3	339,1	284,1	274,9	265,2	1.163,3
M4	379,6	356,1	287,7	279,1	1.302,5
Totais	1288,6	1.254,5	983,3	997,3	4.523,7

$$SQ_M = \frac{1}{16}(918,2^2 + 1.139,7^2 + 1.163,3^2 + 1.302,5^2) - C = 4.738,4405$$

$$SQ_{PARCELAS} = \frac{1}{4}(256,6^2 + 263,7^2 + \dots + 287,7^2 + 279,1^2) - C$$

$$SQ_{PARCELAS} = \frac{1}{4}(1.322.252,29) - 319.747,8389 = 10.815,2336$$

$$SQ_{RES(a)} = SQ_{PARCELAS} - SQ_{BLOCOS} - SQ_M$$

$$SQ_{RES(a)} = 10.815,2337 - 4.986,3105 - 4.738,4405 = 1.090,4827$$

	A1	A2	A3	A4	Totais
M1	200,6	270,8	247,3	199,5	918,2
M2	271,4	296,0	283,0	289,3	1.139,7
M3	287,3	273,7	300,7	301,6	1.163,3
M4	329,0	338,0	309,6	325,9	1.302,5
Totais	1.088,3	1.178,5	1.140,6	1.116,3	4.523,7

$$SQ_A = \frac{1}{16}(1.088,3^2 + 1.178,5^2 + 1.140,6^2 + 1.116,3^2) - C = 274,2355$$

$$SQ_{M,A} = \frac{1}{4}(200,6^2 + 271,4^2 + \dots + 301,6^2 + 325,9^2) - C = 5.995,1587$$

$SQ_{M,A}$: é a soma de quadrados das combinações entre os fatores M e A .

$$SQ_{M \times A} = SQ_{M,A} - SQ_M - SQ_A$$

$$SQ_{M \times A} = 5.995,1587 - 4.738,4405 - 274,2355 = 982,4827$$

$$SQ_{RES(b)} = SQ_{TOTAL} - SQ_{PARCELAS} - SQ_A - SQ_{M \times A}$$

$$SQ_{RES(b)} = 13.252,3511 - 10.815,2337 - 274,2355 - 982,4827 = 1.180,3992$$

Tabela 8.1.4. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	3	4.986,3105	1.662,1034	13,72**
Manejo (M)	3	4.738,4405	1.579,4802	13,04**
Resíduo(a)	9	1.090,4827	121,1647	-----
(Parcelas)	(15)	10.815,2337	-----	-----
Adubação (A)	3	274,2355	91,4118	2,79 <i>ns</i>
Interação M×A	9	982,4827	109,1647	3,33 **
Resíduo (b)	36	1.180,3992	32,78887	-----
Total	63	13.252,3511		

$$F_{3; 9; 0,05} = 3,86 \quad F_{3; 36; 0,05} = 2,87 \quad F_{9; 36; 0,05} = 2,16$$

$$F_{3; 9; 0,01} = 6,99 \quad F_{3; 36; 0,01} = 4,38 \quad F_{9; 36; 0,01} = 2,96$$

Tabela 8.1.5. ANOVA realizada no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	4986,310469	1662,103490	50,69	< 0,0001
M	3	4738,440469	1579,480156	48,17	< 0,0001
Blocos×M	9	1090,482656	121,164740	3,70	0,0023
A	3	274,235469	91,411823	2,79	0,0545
M×A	9	982,482656	109,164740	3,33	0,0047
Resíduo	36	1180,39938	32,78887		
Total	63	13252,35109			

Tabela 8.1.6. ANOVA com as correções no SAS.

Testes de Hipótese, usando a interação Blocos*M como resíduo					
F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	4986,310469	1662,103490	13,72	0,0010
M	3	4738,440469	1579,480156	13,04	0,0013

O programa **Prog24**, em anexo no apêndice B, efetua a análise do ensaio em parcelas subdivididas. O comando “TEST H=Bloco M E=bloco*M;” corrige os valores do teste *F* referentes aos blocos e aos tratamentos principais (*M*), mostrados na tabela 8.1.6, em que os seus quadrados médios devem ser divididos pelo quadrado médio da interação entre os blocos e o fator *M*, isto é, o “RES (a)”.

Considerando que a interação entre os fatores M (métodos de manejo) e A (tipos de adubação) é significativa, não é possível fazer comparações separadamente entre os níveis desses fatores. Logo podem ser feitos os seguintes desdobramentos:

a) Estudo dos tratamentos secundários (adubação) em cada tratamento principal (manejo).

	A1	A2	A3	A4	Totais
M1	200,6	270,8	247,3	199,5	918,2
M2	271,4	296,0	283,0	289,3	1.139,7
M3	287,3	273,7	300,7	301,6	1.163,3
M4	329,0	338,0	309,6	325,9	1.302,5
Totais	1.088,3	1.178,5	1.140,6	1.116,3	4.523,7

$$SQ_{Ad.M1} = \frac{1}{4}(200,6^2 + 270,8^2 + 247,3^2 + 199,5^2) - \frac{1}{16}(918,2)^2 = 939,4325$$

$$SQ_{Ad.M2} = \frac{1}{4}(271,4^2 + 296,0^2 + 283,0^2 + 289,3^2) - \frac{1}{16}(1.139,7)^2 = 82,1069$$

$$SQ_{Ad.M3} = \frac{1}{4}(287,3^2 + 273,7^2 + 300,7^2 + 301,6^2) - \frac{1}{16}(1.163,3)^2 = 129,8269$$

$$SQ_{Ad.M4} = \frac{1}{4}(329,0^2 + 338,0^2 + 309,6^2 + 325,9^2) - \frac{1}{16}(1.302,5)^2 = 105,3519$$

Verificação:

$$SQ_{Ad.M1} + SQ_{Ad.M2} + SQ_{Ad.M3} + SQ_{Ad.M4} = SQ_A + SQ_{A \times M}$$

$$939,4325 + 82,1069 + 129,8269 + 105,3519 = 274,2355 + 982,4827 = 1.256,7182$$

Tabela 8.1.7. ANOVA: adubação dentro dos métodos de manejo.

Contrastes	G.L.	S.Q.	Q.M.	F
Adubo dentro de M1	3	939,4325	313,1441	9,55**
Adubo dentro de M2	3	82,1069	27,3689	0,83 <i>ns</i>
Adubo dentro de M3	3	129,8269	43,2756	1,32 <i>ns</i>
Adubo dentro de M4	3	105,3519	35,1173	1,07 <i>ns</i>
Resíduo (b)	36	1.180,3992	32,7889	

$$F_{3;36;0,05} = 2,87 \quad F_{3;36;0,01} = 4,39$$

A seguir, o desdobramento feito no SAS para adubação dentro dos métodos de manejo:

Adubação dentro dos métodos de manejo					
M	G.L.	S.Q.	Q.M.	F	Pr > F
M1	3	939.432500	313.144167	9.55	< 0.0001
M2	3	82,106875	27,368958	0,83	0,4837
M3	3	129,826875	43,275625	1,32	0,2830
M4	3	105,351875	35,117292	1,07	0,3735

Na análise dos contrastes dos tipos de adubação, dentro dos métodos de manejo, verifica-se que somente há diferença significativa no método de manejo florestal *M1*. Considerar a aplicação do teste de Tukey para os tratamentos de adubação dentro do método de manejo *M1* (*Ad.M1*).

	A1	A2	A3	A4
Totais (M1)	200,6	270,8	247,3	199,5

$$\Delta = q_{4; 36; 0,05} \times \sqrt{\frac{QM_{RES(b)}}{J}} = \frac{3,81 \times \sqrt{32,78887}}{\sqrt{4}} = 10,91$$

$$\overline{M1A2} = \frac{270,8}{4} = 67,70.....a$$

$$\overline{M1A3} = \frac{247,3}{4} = 61,83....a$$

$$\overline{M1A1} = \frac{200,6}{4} = 50,15.....b$$

$$\overline{M1A4} = \frac{199,5}{4} = 49,88.....b$$

Médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem pelo teste de Tukey. A comparação entre as combinações *M1A2* e *M1A3* foi não significativa, assim como apresentaram as maiores produções.

b) Estudo dos tratamentos principais (método de manejo) dentro de cada tratamento secundário (adubação).

	A1	A2	A3	A4	Totais
M1	200,6	270,8	247,3	199,5	918,2
M2	271,4	296,0	283,0	289,3	1.139,7
M3	287,3	273,7	300,7	301,6	1.163,3
M4	329,0	338,0	309,6	325,9	1.302,5
Totais	1.088,3	1.178,5	1.140,6	1.116,3	4.523,7

$$SQ_{Md.A1} = \frac{1}{4}(200,6^2 + 271,4^2 + 287,3^2 + 329,0^2) - \frac{1.088,3^2}{16} = 2.145,3469$$

$$SQ_{Md.A2} = \frac{1}{4}(270,8^2 + 296,0^2 + 273,7^2 + 338,0^2) - \frac{1.178,5^2}{16} = 722,1919$$

$$SQ_{Md.A3} = \frac{1}{4}(247,3^2 + 283,0^2 + 300,7^2 + 309,6^2) - \frac{1.140,6^2}{16} = 569,2125$$

$$SQ_{Md.A4} = \frac{1}{4}(199,5^2 + 289,3^2 + 301,6^2 + 325,9^2) - \frac{1.116,3^2}{16} = 2.284,1719$$

Verificação:

$$SQ_{Md.A1} + SQ_{Md.A2} + SQ_{Md.A3} + SQ_{Md.A4} = SQ_M + SQ_{M \times A}$$

$$SQ_M + SQ_{M \times A} = 4.738,4405 + 982,4827 = 5.720,9232$$

$$SQ_{Md.A1} + SQ_{Md.A2} + SQ_{Md.A3} + SQ_{Md.A4} = 5.720,9232$$

Método de manejo dentro de cada tratamento secundário (adubação)			
A	G.L	S.Q.	Q.M.
A1	3	2145,346875	715,115625
A2	3	722,191875	240,730625
A3	3	569,212500	189,737500
A4	3	2284,171875	761,390625

No estudo abrangendo os tratamentos principais (parcelas (A): manejo florestal), em que cada tratamento secundário (subparcelas (B): adubação) envolve dois resíduos, utiliza-se, pois, o resíduo médio:

$$QM_{RES(MÉDIO)} = \frac{QM_{RES(a)} + (b-1)QM_{RES(b)}}{b}$$

Considerar, ainda, que o número de graus de liberdade (n'_{gl}), associado ao resíduo médio, é calculado, utilizando-se a fórmula de aproximação de Satterthwaite (1946).

$$n'_{gl} = \frac{[QM_{RES(a)} + (b-1)QM_{RES(b)}]^2}{\frac{[QM_{RES(a)}]^2}{(a-1)(J-1)} + \frac{(b-1)^2[QM_{RES(b)}]^2}{a(J-1)(b-1)}}$$

Observação: o valor de n'_{gl} deve ser arredondado para o inteiro mais próximo.

Tal que: $nglRES(a) \leq n'_{gl} \leq [nglRES(a) + nglRES(b)]$

Aplicando-se, no exemplo, tal que b é igual ao número de níveis do fator A (adubo):

$$QM_{RESIDUO}(Médio) = \frac{QM_{RES(a)} + (b-1)QM_{RES(b)}}{b} = \frac{121,1647 + (4-1) \times 32,78887}{4}$$

$$QM_{RESIDUO}(Médio) = 54,8828$$

$$n'_{gl} = \frac{[121,1647 + (4-1)(32,78887)]^2}{\frac{121,1647^2}{9} + \frac{(4-1)^2(32,78887)^2}{36}}$$

$$n'_{gl} = \frac{48.193,9961}{1.631,2094 + 268,7775} \cong 25$$

Verifica-se que a condição $nglRES(a) \leq n'_{gl} \leq [nglRES(a) + nglRES(b)]$ foi atendida:

$$9 < n'_{gl} = 25 < 9 + 36 = 45$$

Tabela 8.1.8. ANOVA: método de manejo dentro dos tipos de adubação.

Contrastes	G.L	S.Q.	Q.M.	F
Manejo dentro de A1	3	2.145,3469	715,1156	13,03**
Manejo dentro de A2	3	722,1919	240,7306	4,39*
Manejo dentro de A3	3	569,2125	189,7375	3,46*
Manejo dentro de A4	3	2.284,1719	761,3906	13,87**
Resíduo médio	25		54,8828	

$$F_{3;25;0,05} = 2,99 \quad F_{3;25;0,01} = 4,68$$

Na análise dos contrastes dos tratamentos de manejo florestal, dentro dos tipos de adubação, verifica-se que há diferença significativa entre os métodos de manejo nos quatro tipos de adubação. Seja, então, considerado o teste de Tukey para $Md.A1$, $Md.A2$, $Md.A3$, $Md.A4$.

$$\Delta = q_{4;25;0,05} \times \frac{\sqrt{RES(MÉDIO)}}{\sqrt{J}} = 3,89 \times \frac{\sqrt{54,8828}}{\sqrt{4}} = 14,41$$

Resultando em:

	A1		A2		A3		A4
M1	50,15 a	M1	67,70 a	M1	61,83 a	M1	49,88 a
M2	67,85 b	M3	68,43 a	M2	70,75 ab	M2	72,33 b
M3	71,83 bc	M2	74,00 ab	M3	75,18 ab	M3	75,40 b
M4	82,25 c	M4	84,50 b	M4	77,40 b	M4	81,48 b

Médias seguidas pela mesma letra não diferem pelo teste de Tukey ao nível de 5% de significância. Portanto o método de manejo M4 apresentou o melhor resultado considerando os quatro tipos de adubação.

8.2. EXPERIMENTOS EM FAIXAS “*SPLIT BLOCK*”

Nos esquemas experimentais em faixas, os tratamentos correspondentes às subunidades, diferentemente dos ensaios em parcelas subdivididas, não são casualizados de forma independente dentro de cada parcela, mas distribuídos com objetivo de formar faixas nas repetições ou blocos.

Esse esquema experimental em faixas, considerando os fatores A e B , compromete a precisão nas estimativas dos efeitos principais de A e B , possibilitando maior precisão para a interação $A \times B$. Por outro lado, em muitas pesquisas experimentais, por questões operacionais, na instalação e na condução do ensaio, ocorrem grandes dificuldades na utilização de parcelas subdivididas, especialmente em ensaios de campo. Nesses ensaios, surge a necessidade de se estabelecer uma restrição dentro de cada bloco, tal que cada nível — tanto do fator A como do fator B —, seja disposto em faixas e em direções transversais de um fator em relação ao outro, caracterizando dois tipos de parcelas, uma para cada fator, ou seja, as combinações dos níveis dos dois fatores representarão as subparcelas.

Muitas vezes, por razões práticas, é recomendável utilizar os experimentos em faixas, como, por exemplo, em um ensaio envolvendo o uso de máquinas na preparação do solo, níveis de irrigação, aplicação mecanizada de adubos, sistemas de exploração florestal, aplicação de inseticidas, entre outros. Ou, ainda, em outro tipo de procedimento, no qual se torna impossível aplicar o tratamento especificado na subparcela. Por outro lado, no intuito de avaliar dois tipos de fatores, existindo a necessidade de distribuir os tratamentos em parcelas relativamente grandes, a utilização de experimentos em faixas deve ser considerada.

Na aplicação de um determinado trato silvicultural, numa área de manejo florestal, por questões de facilidades na gestão da coordenação dos trabalhos, a adoção do experimento em faixas pode tornar mais fácil orientar as equipes de campo, principalmente na execução das tarefas especificadas na administração dos tratamentos.

Considerando o delineamento casualizado em blocos e os fatores A e B , seja considerado o modelo matemático:

$$x_{ikj} = m + \alpha_i + \theta_j + (\alpha\theta)_{ij} + \beta_k + (\beta\theta)_{kj} + (\alpha\beta)_{ik} + (\alpha\theta\beta)_{ikj}$$

x_{ikj} : valor da variável resposta da combinação correspondente ao nível (i) do fator A e ao nível (k) do fator B no bloco (j);

m : média geral;

α_i : efeito do nível (i) do fator A ;

θ_j : considerando o efeito aleatório do j -ésimo bloco, tal que $N(0, \sigma_\theta^2)$;

θ_j : efeito fixo do j -ésimo bloco, tal que $E(\theta_j) = \theta_j$, $E(\theta_j^2) = \theta_j^2$ e $\sum_{j=1}^J \theta_j = 0$;

$(\alpha\theta)_{ij}$: efeito aleatório da interação do i -ésimo nível do fator A com o j -ésimo bloco [erro da parcela (ij)];

β_k : efeito do nível (k) do fator B ;

$(\beta\theta)_{kj}$: efeito aleatório da interação do k -ésimo nível do fator B com o j -ésimo bloco [erro associado da parcela (kj)];

$(\alpha\beta)_{ik}$: efeito da interação do i -ésimo nível do fator A com o k -ésimo nível do fator B ;

$(\alpha\theta\beta)_{ikj}$: efeito aleatório da interação do i -ésimo nível do fator A com o k -ésimo nível do fator B no j -ésimo bloco [erro da parcela (ikj)].

$(\alpha\theta)_{ij}$, $(\beta\theta)_{kj}$ e $(\alpha\theta\beta)_{ikj}$ são os erros experimentais, respectivamente, para os tratamentos A , B e para a interação $A \times B$, normal e independentemente distribuídos, com média zero e variâncias: $\sigma_{\alpha\theta}^2$, $\sigma_{\beta\theta}^2$ e $\sigma^2 = \sigma_{\alpha\theta\beta}^2$.

De acordo com Zimmermann (1973), na tabela 8.2.1, os componentes de variância.

Tabela 8.2.1. Componentes de variância do experimento em faixas.

F.V.	G.L.	E(Q.M.)
Blocos (aleatório)	$J - 1$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + a\sigma_{\beta\theta}^2 + ab\sigma_{\theta}^2$
A	$a - 1$	$\sigma^2 + b\sigma_{\alpha\theta}^2 + \frac{bJ}{a-1} \sum_{i=1}^a \alpha_i^2$
$A \times \text{Blocos: Res(a)}$	$(a-1)(J-1)$	$\sigma^2 + b\sigma_{\alpha\theta}^2$
B	$(b-1)$	$\sigma^2 + a\sigma_{\beta\theta}^2 + \frac{aJ}{b-1} \sum_{k=1}^b \beta_k^2$
$B \times \text{Blocos: Res(b)}$	$(b-1)(J-1)$	$\sigma^2 + a\sigma_{\beta\theta}^2$
$A \times B$	$(a-1)(b-1)$	$\sigma^2 + \frac{J}{(a-1)(b-1)} \sum_{i,k} (\alpha\beta)_{ik}^2$
$A \times B \times \text{Blocos: Res (c)}$	$(a-1)(b-1)(J-1)$	σ^2
Total	$(abJ - 1)$	

Considerar um ensaio, cujo objetivo seja testar quatro métodos de manejo florestal, $M1$, $M2$, $M3$ e $M4$ (Fator A). Nele, existem quatro espécies florestais — $E1$, $E2$, $E3$ e $E4$ (Fator B) —, repetidas em quatro blocos. Observar as seguintes alternativas de análises:

Fatorial		Parcelas subdivididas		Experimento em faixas	
F.V.	G.L.	F.V.	G.L.	F.V.	G.L.
Blocos	3	Blocos	3	Blocos	3
Tratamentos A	3	Tratamentos A	3	Tratamentos A	3
Tratamentos B	3	Resíduo (a)	9	Resíduo (a)	9
Interação $A \times B$	9	(Parcelas)	(15)	Tratamentos B	3
Resíduo	45	Tratamentos B	3	Resíduo (b)	12
Total	63	Interação $A \times B$	9	Interação $A \times B$	9
		Resíduo (b)	36	Resíduo (c)	27
		Total	63	Total	63

Em um experimento, com fatores A e B , que apresentam quatro níveis ($a = 4$ e $b = 4$), em que cada bloco poderia ser dividido em faixas horizontais B_i e faixas verticais A_k . Analogamente, os demais blocos, diferentes entre si, são planejados apenas com referência à casualização.

	A_4	A_2	A_3	A_1
B_3	A_4B_3	A_2B_3	A_3B_3	A_1B_3
B_1	A_4B_1	A_2B_1	A_3B_1	A_1B_1
B_4	A_4B_4	A_2B_4	A_3B_4	A_1B_4
B_2	A_4B_2	A_2B_2	A_3B_2	A_1B_2

O planejamento, no esquema fatorial, nesse caso é o ideal, pois estima os efeitos com maior precisão, porque apresenta um único resíduo para o cálculo do valor de F e para efetuar as comparações múltiplas e, conseqüentemente, maximiza o número de seus graus de liberdade.

A análise do ensaio em parcelas subdivididas é mais simples do que o experimento em faixas, visto que estima com maior precisão, as comparações entre os níveis do fator B e apresenta mais facilidade na interpretação da interação $A \times B$. Os esquemas, em parcelas subdivididas, apresentam dois resíduos, um para comparar as parcelas e outro para comparar as subparcelas. O número de graus de liberdade para as parcelas (fator A) é pequeno, resultando em pouca sensibilidade em sua análise.

O experimento em faixas apresenta três resíduos e, conseqüentemente, ocorrem reduções significativas no número de graus de liberdade para os testes. Os experimentos em faixas somente devem ser usados quando razões operacionais de implantação e de condução justificarem. Piedade (1987) apresenta os procedimentos para efetuar a análise conjunta de grupos de experimentos em faixas.

Exercício 8.2.1. Considerar um ensaio em faixas, cujo o objetivo foi estudar o crescimento de quatro espécies florestais E1, E2, E3 e E4 (Fator A), submetidas a quatro espaçamentos (Fator B), tal que $B1 = 1,5 \text{ m} \times 1,5 \text{ m}$; $B2 = 2 \text{ m} \times 2 \text{ m}$; $B3 = 2,5 \text{ m} \times 2,5 \text{ m}$ e $B4 = 3 \text{ m} \times 3 \text{ m}$. A tabela 8.2.2 apresenta os volumes de madeira, dados simulados, em $\text{m}^3 \text{ ha}^{-1}$, mensurados no final da rotação economicamente aconselhável.

Tabela 8.2.2. Volume de madeira em $\text{m}^3 \text{ ha}^{-1}$.

A	B	Blocos	Prod.	Blocos	Prod.	Blocos	Prod.	Blocos	Prod.
A1	B1	1	152,4	2	153,3	3	103,5	4	117,0
A1	B2	1	190,8	2	204,6	3	153,6	4	162,0
A1	B3	1	173,1	2	188,4	3	142,2	4	145,8
A1	B4	1	157,2	2	146,1	3	99,0	4	121,2
A2	B1	1	187,8	2	243,3	3	158,7	4	122,7
A2	B2	1	201,3	2	246,0	3	148,2	4	181,5
A2	B3	1	209,4	2	230,4	3	144,6	4	158,7
A2	B4	1	224,4	2	201,0	3	154,2	4	180,3

Tabela 8.2.2. Volume de madeira em m³ ha⁻¹ (continuação).

A	B	Blocos	Prod.	Blocos	Prod.	Blocos	Prod.	Blocos	Prod.
A3	B1	1	218,1	2	204,6	3	156,0	4	175,8
A3	B2	1	222,0	2	176,1	3	157,2	4	163,2
A3	B3	1	225,6	2	169,2	3	219,0	4	175,8
A3	B4	1	225,0	2	196,2	3	189,6	4	181,2
A4	B1	1	261,3	2	229,5	3	188,7	4	184,2
A4	B2	1	246,0	2	235,5	3	201,3	4	204,6
A4	B3	1	239,1	2	228,6	3	167,1	4	178,2
A4	B4	1	250,5	2	241,5	3	198,0	4	165,6
			3.384,0		3.294,3		2.580,9		2.617,8

$$C = \frac{G^2}{IJ} = \frac{11.877,0^2}{64} = 2.204.111,3906$$

$$SQ_{TOTAL} = 152,4^2 + 190,8^2 + \dots + 178,2^2 + 165,6^2 - C$$

$$SQ_{TOTAL} = 2.295.684,18 - 2.204.111,3906 = 91.572,7894$$

$$SQ_{BLOCOS} = \frac{1}{16} (3.384,0^2 + 3.294,3^2 + 2.580,9^2 + 2.617,8^2) - C = 34.500,4931$$

	Bloco1	Bloco2	Bloco3	Bloco4	Total
A1	673,5	692,4	498,3	546,0	2.410,2
A2	822,9	920,7	605,7	643,2	2.992,5
A3	890,7	746,1	721,8	696,0	3.054,6
A4	996,9	935,1	755,1	732,6	3.419,7
Total	3.384,0	3.294,3	2.580,9	2.617,8	11.877,0

$$SQ_A = \frac{1}{16} (2.410,2^2 + 2.992,5^2 + 3.054,6^2 + 3.419,7^2) - C = 32.704,2056$$

$$SQ_{A,BLOCOS} = \frac{1}{4} (673,5^2 + 822,9^2 + \dots + 696,0^2 + 732,6^2) - C$$

$$SQ_{A,BLOCOS} = \frac{1}{4} (9.115.393,86) - 2.204.111,3906 = 74.737,0744$$

$SQ_{A,BLOCOS}$: é a soma de quadrados, considerando as combinações dos níveis do fator A e os blocos.

$$SQ_{RES(a)} = SQ_{A \times BLOCOS} = SQ_{A,BLOCOS} - SQ_A - SQ_{BLOCOS}$$

$$SQ_{RES(a)} = SQ_{A \times BLOCOS} = 74.737,0744 - 32.704,2056 - 34.500,4931$$

$$SQ_{RES(a)} = SQ_{A \times BLOCOS} = 7.532,3756$$

	Bloco1	Bloco2	Bloco3	Bloco4	Total
B1	819,6	830,7	606,9	599,7	2.856,9
B2	860,1	862,2	660,3	711,3	3.093,9
B3	847,2	816,6	672,9	658,5	2.995,2
B4	857,1	784,8	640,8	648,3	2.931,0
Total	3.384,0	3.294,3	2.580,9	2.617,8	11.877,0

$$SQ_B = \frac{1}{16}(2.856,9^2 + 3.093,9^2 + 2.995,2^2 + 2.931,0^2) - C = 1.893,5381$$

$$SQ_{B,BLOCOS} = \frac{1}{4}(819,6^2 + 860,1^2 + \dots + 658,5^2 + 648,3^2) - C$$

$SQ_{B,BLOCOS}$: é a soma de quadrados, considerando as combinações dos níveis do fator B e os blocos.

$$SQ_{B,BLOCOS} = \frac{1}{4} \times 8.967.328,02 - C = 37.720,6144$$

$$SQ_{RES(b)} = SQ_{B \times BLOCOS} = SQ_{B,BLOCOS} - SQ_B - SQ_{BLOCOS}$$

$$SQ_{RES(b)} = SQ_{B \times BLOCOS} = 37.720,6144 - 1.893,5382 - 34.500,4931$$

$$SQ_{RES(b)} = SQ_{B \times BLOCOS} = 1.326,5831$$

	A1	A2	A3	A4	Total
B1	526,2	712,5	754,5	863,7	2.856,9
B2	711,0	777,0	718,5	887,4	3.093,9
B3	649,5	743,1	789,6	813,0	2.995,2
B4	523,5	759,9	792,0	855,6	2.931,0
Total	2.410,2	2.992,5	3.054,6	3.419,7	11.877,0

$$SQ_{A,B} = \frac{1}{4}(526,2^2 + 711,0^2 + \dots + 813,0^2 + 855,6^2) - C$$

$SQ_{A,B}$: é a soma de quadrados das combinações dos níveis do fator A e dos níveis do fator B .

$$SQ_{A,B} = \frac{1}{4}(8.982.062,28) - C = 41.404,1794$$

$$SQ_{A \times B} = SQ_{A,B} - SQ_A - SQ_B = 41.404,1795 - 32.704,2057 - 1.893,5382 = 6.806,4356$$

$$SQ_{RES(c)} = SQ_{TOTAL} - SQ_{BLOCOS} - SQ_A - SQ_{RES(a)} - SQ_B - SQ_{RES(b)} - SQ_{A \times B}$$

$$SQ_{RES(c)} = 6.809,1581$$

Tabela 8.2.3. ANOVA.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos	3	34.500,4931	11.500,1644	13,74 **
Tratamentos (A)	3	32.704,2056	10.901,4019	13,03 **
Resíduo (a)	9	7.532,3756	836,9306	
Tratamentos (B)	3	1.893,5381	631,1794	4,28*
Resíduo (b)	9	1.326,5831	147,3981	
Interação A×B	9	6.806,4356	756,2706	3,00*
Resíduo (c)	27	6.809,1581	252,1910	
Total	63	91.572,7892		

Tabela 8.2.4. ANOVA no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	34500,49313	11500,16438	45,60	< 0,0001
Trat. (A)	3	32704,20562	10901,40187	43,23	< 0,0001
Blocos×A: RES (a)	9	7532,37563	836,93063	3,32	0,0075
Trat. (B)	3	1893,53813	631,17938	2,50	0,0806
Blocos×B: RES (b)	9	1326,58312	147,39812	0,58	0,7981
A×B	9	6806,43562	756,27062	3,00	0,0130
RES (c)	27	6809,15812	252,19104		
Total	63	91572,78938			

Testes de Hipóteses, usando o Q.M. de Blocos*A como resíduo					
F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos	3	34500,49313	11500,16438	13,74	0,0010
Trat.(A)	3	32704,20562	10901,40187	13,03	0,0013

Teste de Hipótese, usando como Q.M. de Blocos*B como resíduo					
F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Trat. (B)	3	1893,538125	631,179375	4,28	0,0389

$$F_{3; 27; 0,05} = 2,96 \quad F_{3; 27; 0,01} = 4,60$$

O programa **Prog25**, encontrado no apêndice B, efetua a análise do delineamento em faixa e — por meio dos comandos: “TEST H=Bloco A E=Bloco*A;” e “TEST H=B E=Bloco*B;” —, corrige os valores do teste F referentes aos blocos e aos fatores A e B , usando os resíduos adequados em cada caso, isto é, os quadrados médios dos blocos e do fator A devem ser divididos pelo RES (a), o quadrado médio do fator B deve ser dividido pelo RES(b) e a interação $A \times B$ deve ser dividida pelo RES(c).

Na existência de significância da interação entre os fatores A e B , deverão ser efetuados os seguintes desdobramentos:

a) Estudo dos tratamentos B (espaçamentos) em cada tratamento A (espécie).

	A1	A2	A3	A4	Total
B1	526,2	712,5	754,5	863,7	2.856,9
B2	711,0	777,0	718,5	887,4	3.093,9
B3	649,5	743,1	789,6	813,0	2.995,2
B4	523,5	759,9	792,0	855,6	2.931,0
Total	2.410,2	2.992,5	3.054,6	3.419,7	11.877,0

$$SQ_{Bd..A1} = \frac{1}{4}(526,2^2 + 711,0^2 + 649,5^2 + 523,5^2) - \frac{2.410,2^2}{16} = 6.510,9825$$

$$SQ_{Bd..A2} = \frac{1}{4}(712,5^2 + 777,0^2 + 743,1^2 + 759,9^2) - \frac{2.992,5^2}{16} = 566,7018$$

$$SQ_{Bd..A3} = \frac{1}{4}(754,5^2 + 718,5^2 + 789,6^2 + 792,0^2) - \frac{3.054,6^2}{16} = 899,8425$$

$$SQ_{Bd..A4} = \frac{1}{4}(863,7^2 + 887,4^2 + 813,0^2 + 855,6^2) - \frac{3.419,7^2}{16} = 722,4469$$

$$\text{Verificação: } SQ_{Bd..A1} + SQ_{Bd..A2} + SQ_{Bd..A3} + SQ_{Bd..A4} = SQ_B + SQ_{A \times B}$$

$$6.510,9825 + 566,7018 + 899,8425 + 722,4469 = 8.699,9737$$

$$SQ_B + SQ_{A \times B} = 1.893,5381 + 6.806,4356 = 8.699,9737$$

O estudo abrangendo os tratamentos do fator B (espaçamento) em cada tratamento do fator A (espécie), envolve dois resíduos, então se utiliza o resíduo médio, tal que a é igual ao número de níveis do fator A .

$$QM_{RES. \text{ MÉDIO}(b,c)} = \frac{QM_{RES(b)} + (a-1)QM_{RES(c)}}{a}$$

$$QM_{RES. \text{ MÉDIO}(b,c)} = \frac{147,3981 + 3 \times 252,1910}{4} = 225,9928$$

$$n'_{gl} = \frac{[QM_{RES(b)} + (a-1)QM_{RES(c)}]^2}{\frac{[QM_{RES(b)}]^2}{nglRES(b)} + \frac{(a-1)^2[QM_{RES(c)}]^2}{nglRES(c)}}$$

$$n'_{gl} = \frac{(147,3981 + 3 \times 252,1910)^2}{\frac{147,3981^2}{9} + \frac{3^2 \times 252,1910^2}{27}}$$

$$n'_{gl} = \frac{817.163,7496}{2.414,0222 + 21.200,1002} \cong 35$$

Tabela 8.2.5. ANOVA para espaçamento (B) em cada espécie (A).

Contrastes	G.L.	S.Q.	Q.M.	F
B dentro de A1 (E1)	3	6.510,9825	2.170,3275	9,60**
B dentro de A2 (E2)	3	566,7019	188,9006	0,84 ns
B dentro de A3 (E3)	3	899,8425	299,9475	1,33 ns
B dentro de A4 (E4)	3	722,4469	240,8156	1,07 ns
Resíduo médio	35		225,9928	

$$F_{3;35;0,05} = 3,28 \quad F_{3;35;0,01} = 4,41$$

Tratamentos B (espaçamentos) em cada tratamento A (espécie).			
A	G.L.	S.Q.	Q.M.
A1	3	6510,982500	2170,327500
A2	3	566,701875	188,900625
A3	3	899,842500	299,947500
A4	3	722,446875	240,815625

Na análise dos contrastes dos espaçamentos, dentro de cada fator A (espécie), verifica-se que há apenas diferença significativa entre os quatro espaçamentos dentro de A1 (E1). Assim, deve-se aplicar o teste de Tukey para B.d.A1, apresentado a seguir:

Teste de Tukey dos tratamentos B (espaçamentos) para A1 (Espécie E1).

$$\Delta = q_{4;35;0,05} \times \frac{\sqrt{RESÍDUO \text{ MÉDIO}}}{\sqrt{J}} = 3,82 \times \frac{\sqrt{225,9928}}{\sqrt{4}} = 28,71$$

Logo:

Espaçamento (B)	A1 (Espécie E1)
B4	130,88 a
B1	131,55 a
B3	162,38 b
B2	177,75 b

Médias seguidas pela mesma letra não diferem, pelo teste de Tukey, ao nível de 5% de significância. A espécie E1 (A1) apresentou o melhor resultado, considerando os tipos de espaçamentos $B1 = 1,5 \text{ m} \times 1,5 \text{ m}$ e $B2 = 2 \text{ m} \times 2 \text{ m}$.

b) Estudo dos tratamentos A (espécies) em cada tratamento B (espaçamento).

$$SQ_{Ad.B1} = \frac{1}{4}(526,2^2 + 712,5^2 + 754,5^2 + 863,7^2) - \frac{2.856,9^2}{16} = 14.830,3068$$

$$SQ_{Ad.B2} = \frac{1}{4}(711,0^2 + 777,0^2 + 718,5^2 + 887,4^2) - \frac{3.093,9^2}{16} = 4.979,1769$$

$$SQ_{Ad.B3} = \frac{1}{4}(649,5^2 + 743,1^2 + 789,6^2 + 813,0^2) - \frac{2.995,2^2}{16} = 3.919,8150$$

$$SQ_{Ad.B4} = \frac{1}{4}(523,5^2 + 759,9^2 + 792,0^2 + 855,6^2) - \frac{2.931,0^2}{16} = 15.781,3425$$

Verificação: $SQ_{Ad.B1} + SQ_{Ad.B2} + SQ_{Ad.B3} + SQ_{Ad.B4} = SQ_A + SQ_{A \times B}$

$$14.830,3068 + 4.979,1769 + 3.919,8150 + 15.781,3425 = 39.510,6412$$

$$SQ_A + SQ_{A \times B} = 32.704,2056 + 6.806,4356 = 39.510,6412$$

O estudo, abrangendo os tratamentos do fator A (espécies) em cada tratamento do fator B (espaçamento), envolve dois resíduos. Nesse caso, utiliza-se o resíduo médio, tal que b é igual ao número de níveis do fator B .

$$QM_{RES.MÉDIO(a,c)} = \frac{QM_{RES(a)} + (b-1)QM_{RES(c)}}{b}$$

$$QM_{RES.MÉDIO(a,c)} = \frac{836,9306 + 3 \times 252,1910}{4} = 398,3759$$

$$n''_{gl} = \frac{[QM_{RES(a)} + (b-1)QM_{RES(c)}]^2}{\frac{[QM_{RES(a)}]^2}{n_{gl}RES(a)} + \frac{(b-1)^2[QM_{RES(c)}]^2}{n_{gl}RES(c)}}$$

$$n''_{gl} = \frac{(836,9306 + 3 \times 252,1910)^2}{\frac{836,9306^2}{9} + \frac{3^2 \times 252,1910^2}{27}}$$

$$n''_{gl} = \frac{2.539.253,7232}{77.828,0921 + 21.200,1002} \cong 26$$

Tabela 8.2.6. ANOVA para espécie (*A*) em cada espaçamento (*B*).

Contrastes	G.L.	S.Q.	Q.M.	F
<i>A</i> dentro de <i>B</i> 1	3	14.830,3068	4.943,4356	12,41**
<i>A</i> dentro de <i>B</i> 2	3	4.979,1769	1.659,7256	4,17*
<i>A</i> dentro de <i>B</i> 3	3	3.919,8150	1.306,6050	3,28*
<i>A</i> dentro de <i>B</i> 4	3	15.781,3425	5.260,4475	13,20**
Resíduo médio	26		398,3759	

$$F_{3; 26; 0,05} = 2,98 \quad F_{3; 26; 0,01} = 4,64$$

Estudo dos tratamentos <i>A</i> (espécie) em cada tratamento <i>B</i> (espaçamento)			
B	G.L.	S.Q.	Q.M.
B1	3	14830,000000	4943,435625
B2	3	4979,176875	1659,725625
B3	3	3919,815000	1306,605000
B4	3	15781,000000	5260,447500

Na análise dos contrastes dos tratamentos do fator *A* (espécies), dentro de cada espaçamento (Fator *B*), verifica-se que há diferença significativa dos quatro tratamentos do fator *A* dentro dos quatro espaçamentos, sugerindo o teste de Tukey para *Ad.B1*, *Ad.B2*, *Ad.B3*, *Ad.B4*, a fim de avaliar suas diferenças.

Teste de Tukey dos tratamentos *A* (espécies) em cada tratamento *B* (espaçamento).

$$\Delta = q_{4; 26; 0,05} \times \frac{\sqrt{RESÍDUO_{MÉDIO}}}{\sqrt{J}} = 3,88 \times \frac{\sqrt{398,3759}}{\sqrt{4}} = 38,72$$

	Médias						
	<i>B</i> 1		<i>B</i> 2		<i>B</i> 3		<i>B</i> 4
<i>A</i> 1	131,55 a	<i>A</i> 1	177,75 a	<i>A</i> 1	162,38 a	<i>A</i> 1	130,88 a
<i>A</i> 2	178,13 b	<i>A</i> 3	179,63 a	<i>A</i> 2	185,78 ab	<i>A</i> 2	189,98 b
<i>A</i> 3	188,63 b	<i>A</i> 2	194,25 ab	<i>A</i> 3	197,40 ab	<i>A</i> 3	198,00 b
<i>A</i> 4	215,93 b	<i>A</i> 4	221,85 b	<i>A</i> 4	203,25 b	<i>A</i> 4	213,90 b

Médias seguidas pela mesma letra não diferem, pelo teste de Tukey, ao nível de 5% de significância. A combinação A4B2 (E4 e B2 = 2 m × 2 m) mostrou a maior produção.

8.3. EXPERIMENTOS COM PARCELAS SUBDIVIDIDAS NO TEMPO (*SPLIT PLOT IN TIME*)

O objetivo do experimento com parcelas subdivididas no tempo é estimar o efeito global dos níveis do fator A , ao longo de um período estabelecido, ou seja, estudar a variação dos níveis ao longo do tempo. É importante perceber que a caracterização das subparcelas (Fator B : ocasiões) denota uma sistematização, haja vista a impossibilidade de casualizar as épocas dentro das parcelas. As parcelas podem ser estruturadas em qualquer delineamento, contudo o mais usual é distribuí-las inteiramente ao acaso ou com casualização em blocos.

Por outro lado, como nos verdadeiros arranjos de parcelas subdivididas, as subparcelas devem ser distribuídas aleatoriamente dentro das parcelas. Alguns autores sugerem, como Steel e Torrie (1960), no caso do delineamento casualizado em blocos, isolar a interação entre os tratamentos secundários e blocos do resíduo (b), tal como a análise de variância proposta do exemplo a seguir:

Considerar um experimento com parcelas permanentes de competição de I espécies florestais, em t ocasiões, distribuídas em J blocos.

F.V.	G.L.
Blocos	$J - 1$
Espécies (A)	$I - 1$
Resíduo (a)	$(I - 1)(J - 1)$
Parcelas	$(IJ - 1)$
Ocasões (B)	$t - 1$
Espécies $\times$ ocasiões (A $\times$ B)	$(I - 1)(t - 1)$
Ocasões $\times$ Blocos (B $\times$ Blocos)	$(t - 1)(J - 1)$
Resíduo (b)	$(I - 1)(t - 1)(J - 1)$
Total	$tIJ - 1$

$$SQ_{OCASI\tilde{O}ES \times BLOCOS} = SQ_{OCASI\tilde{O}ES, BLOCOS} - SQ_{OCASI\tilde{O}ES} - SQ_{BLOCOS}$$

$SQ_{OCASI\tilde{O}ES, BLOCOS}$: é a soma de quadrados das combinações entre as ocasiões e os blocos obtida a partir do quadro auxiliar que relaciona ocasiões e blocos.

No manejo florestal é importante conhecer o perfil de crescimento das espécies para estabelecer os tratos silviculturais, assim como a determinação do ciclo de corte ideal ou rotação economicamente aconselhável. Nos ensaios com culturas perenes é fundamental monitorar os ciclos de produção, visando estabelecer os tipos de tratamentos adequados que maximizem a produção. Nesses tipos de experimentos, o tempo, avaliado em dias, meses, anos etc, é considerado um fator secundário e analisado na forma usual de parcelas subdivididas.

O esquema do experimento de parcelas subdivididas no tempo, para efetuar medidas repetidas em ocasiões sucessivas, é caracterizado quando se aplica às parcelas os níveis de um único fator A . Entretanto, é admitido que essas medidas, quando tomadas em ocasiões distintas, tenham variâncias homogêneas e sejam igualmente correlacionadas, ou seja, a aplicação da análise univariada de parcelas subdividida exige que a matriz de covariância, considerando as ocasiões, seja uniforme. Contudo, quando a condição de uniformidade da matriz não é atendida, a estrutura de medidas repetidas no tempo passa a exigir a aplicação da metodologia de análise multivariada, denominada

de análise de perfil, que adota uma hipótese mais geral sobre a estrutura de covariâncias entre as medidas repetidas tomadas na mesma parcela.

O inventário florestal contínuo utilizando parcelas permanentes, enquadra-se, no contexto da análise estatística, no grupo dos denominados Experimentos com Medidas Repetidas (EMR). Nesse caso, a repetição se dá no tempo. Os EMR constituem um problema de população multivariada, em que as variáveis aleatórias — como, por exemplo, o volume de madeira em p -ocasiões — possuem uma distribuição de probabilidade com vetor de médias $\mu' = [\mu_1, \mu_2, \dots, \mu_p]$ e matriz de covariâncias Σ .

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

O procedimento a ser adotado deverá selecionar e medir as parcelas na primeira ocasião, tornando-as permanentes a serem remedidas nas ocasiões seguintes. Nesse procedimento ocorrerão correlações entre as medições nas diversas ocasiões, significando, para a Ciência Florestal, importante mecanismo para a interpretação do comportamento dos parâmetros envolvidos na dinâmica da floresta.

A aplicação da análise univariada pelo método de parcelas subdivididas, em experimentos de medidas repetidas no tempo, somente é válida sob a pressuposição de uniformidade da matriz de covariância. A matriz uniforme ocorrerá quando as ocasiões apresentarem homogeneidade de variâncias e forem igualmente correlacionadas (ρ), ou seja:

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \cdots & \cdots & \cdots & \cdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

Assim, σ^2 é a variância comum entre as unidades dentro das ocasiões e ρ é a correlação comum entre as ocasiões combinadas aos pares. Por outro lado (no caso de não ser atendida a pressuposição exposta e de haver o interesse em analisar, por meio da modelagem univariada do delineamento em parcelas subdivididas) a correção ε_{gg} de Geisser e Greenhouse (1958) ou a correção ε_{hf} de Huynh e Heldt (1976) podem ser usadas para corrigir os graus de liberdade do teste F , para testar a igualdade dos efeitos de ocasiões e de ocorrência de paralelismo dos perfis de crescimento.

Portanto, existem dois métodos para testar as hipóteses sobre o efeito dos tratamentos e a correspondente tendência de crescimento ao longo do tempo. O primeiro é o procedimento univariado, que trata as observações medidas no tempo como se fossem originadas de subdivisões das unidades experimentais (*split-plot*), sendo analisado conforme o delineamento em parcelas subdivididas. O segundo método, denominado de análise de perfil, *profile analysis*, adota um procedimento multivariado, em outras palavras, considera as observações repetidas sobre cada unidade experimental como um vetor de respostas.

Queiroz (2021) descreve o método de Wilks para testar se a matriz de covariância entre as ocasiões é uniforme; o critério de Mauchly (1940) para verificar a ocorrência de esfericidade da matriz; e as correções de Geisser e Greenhouse (1958) e de Huynh e Heldt (1976) para corrigir os graus de liberdade do teste F , a fim de testar a igualdade dos efeitos de ocasiões e a ocorrência de paralelismo dos perfis de crescimento, caso haja a opção pela escolha da análise univariada do delineamento em parcelas divididas.

Queiroz (2021) aborda e apresenta exemplos da aplicação de análise do perfil de crescimento das espécies em inventários florestais, com parcelas permanentes e expõe um exemplo de aplicação, considerando os perfis de crescimento de quatro espécies florestais, conforme os seguintes temas: teste de esfericidade de matriz de covariância pelo critério de Mauchly; cálculo dos valores das correções de Greenhouse-Geisser Epsilon (ϵ_{gg}) e de Huynh-Feldt Epsilon (ϵ_{hf}); teste de homogeneidade de matrizes de covariâncias; testes de multinormalidade de assimetria e de curtose de Mardia; teste de normalidade multivariada de Henze-Zirkler; análise de agrupamento pelo método de Ward; testes de paralelismo e de coincidência de perfis de crescimento pela hipótese linear geral $H : C\beta U = \phi$; análise de coincidência de perfis pelo teste “ t ” de Student; determinação do ciclo de corte economicamente aconselhável; análise de resíduos; teste de validade de Graybill; e teste da razão da máxima verossimilhança para verificar a igualdade de parâmetros e identidade de modelos de regressão não linear.

O exemplo abaixo (dados simulados), apresentado por Queiroz (2021), considera os vetores de médias para a variável volume de madeira de quatro espécies ($E1$, $E2$, $E3$ e $E4$), respectivamente nas ocasiões 0, 5, 10, 15 e 20 anos.

$$\begin{aligned}\bar{y}'_{v1} &= [9,68449667 \quad 14,4152667 \quad 17,0891900 \quad 19,4454800 \quad 21,0381800] \\ \bar{y}'_{v2} &= [9,74305333 \quad 14,4460933 \quad 17,1927833 \quad 19,5607867 \quad 21,1482133] \\ \bar{y}'_{v3} &= [9,67315333 \quad 13,1066700 \quad 16,2532333 \quad 17,7585900 \quad 19,3376867] \\ \bar{y}'_{v4} &= [9,72483333 \quad 13,2058667 \quad 16,5078900 \quad 18,1222500 \quad 19,6947067]\end{aligned}$$

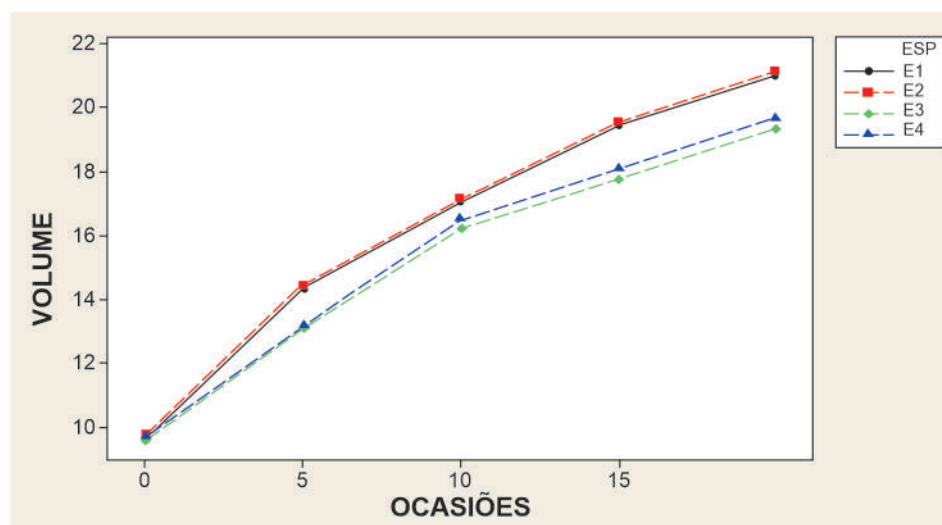


Figura 8.3.1. Perfis de crescimento das espécies $E1$, $E2$, $E3$ e $E4$.

A tabela 8.3.1 mostra que a matriz de covariância, considerando as cinco ocasiões, não é esférica, ou seja, no caso presente ela não é uniforme, então não é adequado usar o método de parcelas subdivididas. Por outro lado, caso haja a opção pela escolha da análise univariada do delineamento em parcelas subdivididas, há os valores das correções de Geisser e Greenhouse (1958) e de Huynh e Feldt (1976) para corrigir os graus de liberdade do teste F , para a igualdade de efeitos de ocasiões e de ocorrência de paralelismo de perfis de crescimento. Atualmente, entretanto, devido à disponibilidade de vários aplicativos computacionais, o mais adequado e preciso no caso da matriz de covariância não ser uniforme, é efetuar a análise de inventário florestal contínuo pela análise multivariada.

Tabela 8.3.1. Teste de esfericidade pelo critério de Mauchly e os valores das correções de Greenhouse-Geisser Epsilon (ϵ_{gg}) e de Huynh-Feldt Epsilon (ϵ_{hf}).

Variáveis	GL	Critério de Mauchly	Qui-Quadrado	Pr > χ^2
Variáveis Transformadas	9	0,0135191	492,40987	< 0,0001
Componentes Ortogonais	9	0,0135191	492,40987	< 0,0001
Greenhouse-Geisser Epsilon	$\hat{\epsilon}_{gg} = 0,3336$			
Huynh-Feldt Epsilon	$\hat{\epsilon}_{hf} = 0,3448$			

No exemplo apresentado em Queiroz (2021), o teste de ocorrência de interação Ocasão×Espécie foi significativo, tabela 8.3.2, isto é, existe interação entre os perfis de crescimento das espécies $E1$, $E2$, $E3$ e $E4$.

Tabela 8.3.2. Teste da interação Ocasão×Espécie para $E1$, $E2$, $E3$ e $E4$.

Testes e aproximações das estatísticas F para a hipótese de não existência de efeito da interação Ocasão×Espécie (S = 3; M = 0; N* = 55,5)					
Estatísticas	Valores	Valor F	Num. GL	Den. GL	Pr > F
Wilks Lambda	0,01633596	93,16	12	299,26	< 0,0001
Pillai Trace	1,20021028	19,17	12	345,00	< 0,0001
Hotelling-Lawley Trace	46,95965330	438,91	12	193,48	< 0,0001
Roy Greatest Root	46,67569000	1.341,93	4	115,00	< 0,0001

Os testes de Wilks Lambda, Pillai Trace, Hotelling-Lawley Trace e de Roy Greatest Root apresentaram resultados significativos (p -valor < 0,0001) para a interação ocasião×espécie, mostrando que existem pelo menos duas espécies que apresentam perfis de crescimento não paralelos. Essas estatísticas estão transformadas na F de Snedecor.

Desse modo, como a interação foi significativa, considerando os perfis de crescimento das quatro espécies florestais, o próximo passo será verificar quais perfis de crescimento são paralelos ou coincidentes. Pode-se realizar um estudo preliminar testando, por exemplo, cada ocasião separadamente pelo teste F , ou aplicar um teste de comparação múltipla, como o teste de Tukey. Queiroz (2021) optou pelo uso da técnica de análise multivariada de agrupamento de Ward. Nele, os grupos foram formados utilizando as ocasiões como variáveis. A figura 8.3.2 apresenta o dendrograma ou fenograma para verificar preliminarmente se os perfis de crescimento das espécies são paralelos. O dendrograma recomenda testar se os perfis de crescimento das espécies $E1$ e $E2$ são paralelos, assim como entre as espécies $E3$ e $E4$.

A partir do dendrograma ou fenograma, resultado da análise de agrupamento, foi realizado o teste para verificar se os perfis de crescimentos das espécies *E1* e *E2* são paralelos, assim como para as espécies *E3* e *E4*.

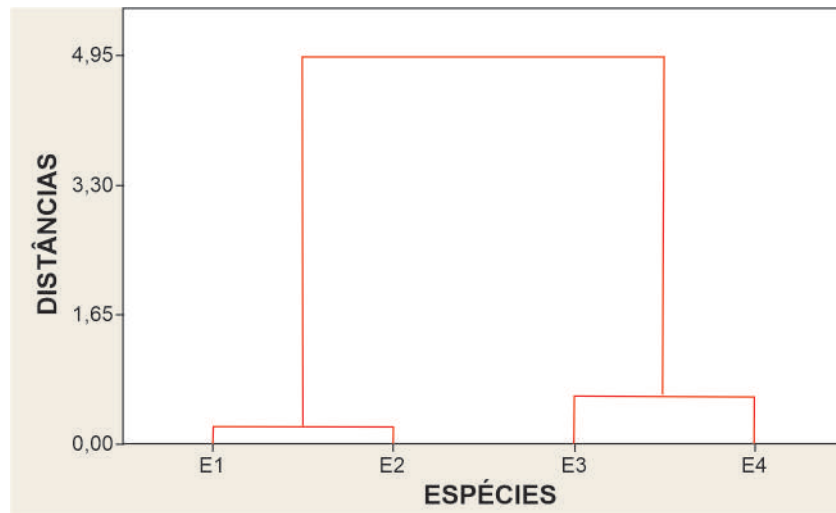


Figura 8.3.2. Dendrograma das quatro espécies florestais.

Os resultados dos testes de Wilks Lambda, Pillai Trace, Hotelling-Lawley Trace e de Roy Greatest Root deram não significantes (p -valor = 0,0904), denotando não existir interação entre os perfis de crescimento das espécies *E1* e *E2*, isto é, os seus perfis são paralelos. O próximo passo será testar se os perfis de crescimento de *E1* e *E2* são coincidentes.

No caso das espécies *E3* e *E4*, os testes indicam existir interação entre os seus perfis de crescimento, ou seja, os seus perfis não são paralelos, então as equações de crescimento para essas espécies devem ser obtidas separadamente.

Os testes de coincidência dos perfis de crescimento de *E1* e *E2*, realizados pelo teste “ t ” de Student, denotaram que os perfis de *E1* e *E2* são coincidentes.

Uma vez observada coincidência entre os perfis de crescimento das espécies *E1* e *E2*, foi obtida uma única equação para ambas as espécies (V_{E12}), a fim de estimar o ciclo de corte do grupo formado pelas espécies (*E1*, *E2*).

Segue a equação de Mitscherlich (V) e a fórmula, de acordo com Queiroz (2021), para obter o ciclo de corte economicamente aconselhável (t_{oti}).

$$V = \beta_0 [1 - 10^{-\beta_1(t + \beta_2)}] \quad t_{oti} = \frac{1}{\hat{\beta}_1} \log_{10} [2,302585 R \hat{\beta}_0 \hat{\beta}_1] - \hat{\beta}_2$$

Na análise foi utilizada uma razão de custo $R = 3$ que é o quociente entre a receita obtida pela venda da madeira e o custo do manejo florestal (por ha). Na tabela 8.3.3. estão apresentadas as estimativas dos parâmetros e a análise de variância.

$$V_{E12} = 24,5911 [1 - 10^{-0,0310 (t + 7,0978)}]$$

$$t_{Oti(E12)} = \left(\frac{1}{0,0310} \right) \log(2,302585 \times 3 \times 24,5911 \times 0,0310) - 7,0978 = 16,2 \text{ anos}$$

$$V_{oti(E12)} = 24,5911 [1 - 10^{-0,0310 (16,2 + 7,0978)}] = 19,929 \text{ m}^3 / \text{ha}$$

$$RQMR_{E12} = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N}} = \sqrt{\frac{38,0422}{2 \times 150}} = 0,3561$$

$$CV_{E12} = \frac{RQMR}{\bar{Y}_{E12}} \times 100 = \frac{0,3561}{16,3764} \times 100 = 2,17\%$$

Tabela 8.3.3. Análise de regressão do grupo formado (E1, E2).

Parâmetros	Estimativas	Erro Padrão Aprox.	Limites de Confiança (95%)	
$\hat{\beta}_0$	24,5911	0,210500	24,1769	25,0053
$\hat{\beta}_1$	0,0310	0,000824	0,0294	0,0326
$\hat{\beta}_2$	7,0978	0,140600	6,8212	7,3745
Fonte	GL	Soma de Quadrados	Quadrado Médio	Valor F
Erro	297	38,0422	0,1281	221.975,0
Teste de validade de Graybill F_G				
Fonte	GL	Quadrado Médio	Valor F_G	Pr > F
Numerador	2	0,14828	1,17	0,3116
Denominador	298	0,12666		
Coeficiente de correlação de Pearson, N = 300, Prob > r Sob H0: Rho=0				
$r_{E12} = 0,99609 (< 0,0001)$			$r_{E12}^2 = 0,9922$	

Pinto (2008), usando a análise univariada de experimentos em parcelas subdivididas, analisou a dinâmica de uma floresta tropical de terra firme, que foi submetida a uma exploração seletiva de madeira, sob diferentes intensidades de corte. Esse estudo foi executado na Estação ZF-2 do INPA e o monitoramento foi realizado em parcelas permanentes. O experimento cobriu uma área de 72 hectares planejado na forma de um delineamento casualizado em blocos, quer dizer, foram apresentados 3 blocos de 24 hectares (400 m x 600 m), 3 tratamentos (intensidades de corte) e uma testemunha de 4 ha (200 m x 200 m). Nesse estudo, foram analisadas 12 parcelas permanentes de 1 ha (100 m x 100 m) instaladas no centro dos tratamentos, tal que todos os indivíduos arbóreos com DAP ≥ 10 cm foram mensurados. Os incrementos (corrente anual: ICA e médio anual: IMA) foram estimados para o período observado de 1990 a 2005.

De acordo com Pinto (2008), as análises estatísticas das variáveis área basal, volume e biomassa, indicaram que o incremento corrente anual (ICA) e o incremento médio anual (IMA), agregando todas as espécies, variaram significativamente com o passar do tempo. Os valores de F foram corrigidos, usando os fatores de correção Greenhouse-Geisser (G-G) ou Huynh-Feldt (H-F) para inferir sobre cada fonte de variação experimental. Na tabela 8.3.4 estão apresentados os valores das correções GG e HF para as 18 variáveis estudadas pelo autor.

Pinto (2008) utilizou, como critério de intervenção na floresta — ou seja, para definir o ciclo do corte ideal —, o ponto ótimo definido quando a curva de incremento corrente anual (ICA) cruzou com a curva do incremento médio anual (IMA), indicando o momento ideal para intervir novamente na floresta. Concluiu que o comportamento linear do incremento médio anual, dentro do período estudado, denotou que o tempo necessário para que as florestas remanescentes apresentem de volta os estoques em volume comercial, semelhante aos estoques da floresta original, deveria ser respectivamente de 30, 52 e 58 anos para os tratamentos T2 (intensidade intermediária de exploração), T3 (intensidade pesada de exploração) e T1 (intensidade leve de exploração).

Tabela 8.3.4. Valores das correções (GG) e (HF) obtidas por Pinto (2008).

Variáveis	Greenhouse-Geisser	Huynh-Feldt
1 - Número de árvores medidas	0,1424	0,25851
2 - Número de árvores comerciais	0,1493	0,3063
3 - ICA para área basal	0,4925	1,0000
4 - ICA para área basal (espécies comerciais)	0,3693	0,7130
5 - IMA para a área basal	0,1635	0,2669
6 - IMA para a área basal (espécies comerciais)	0,0952	0,1459
7- Volume total	0,1321	0,2550
8 - Volume árvores comerciais	0,1219	0,2269
9 - ICA volume total	0,4785	1,0000
10 - ICA volume (espécies comerciais)	0,3698	0,7142
11 - IMA volume total	0,1608	0,2619
12 - IMA volume total para ocasião de 2 anos	0,2204	0,3462
13 - IMA volume total para ocasião de 3 anos	0,3410	0,5413
14 - IMA volume total para ocasião de 4 anos	0,5742	0,8596
15 - IMA volume total para ocasião de 5 anos	0,5895	0,8904
16 - IMA volume espécies comerciais	0,0953	0,1461
17 - IMA em carbono todas as espécies	0,1132	0,1896
18 - IMA em carbono das espécies comerciais	0,1101	0,1825

8.4. BIBLIOGRAFIA

COCHRAN, W. G.; COX, G. M. **Experimental designs**. 2ª ed. New York: John Wiley, 1957. 611 p.

GEISSER, S.; GREENHOUSE, S. W. An extension of Box's results on the use of F distribution in multivariate analysis. **Ann. Math. Statist.**, 29, 855-891, 1958.

HUYNH, H.; FELDT, L.S. Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. **J. Ed. Statist.**, 1, 69-82, 1976.

MAUCHLY, J. W. Significance Test for Sphericity of a Normal n-Variate Distribution. **The Annals of Mathematical Statistics**. Baltimore, v.11, n.6, p. 204-209, 1940.

- PIEDADE, S. M. S. **Análise de grupos de experimentos em faixas**. Dissertação de mestrado. Pós-graduação em estatística e experimentação agronômica. Piracicaba: ESAQ-USP, 1987. 162 p.
- PINTO, A. C. M. **Dinâmica de uma floresta de terra firme manejada experimentalmente na região de Manaus (AM)**. 182 p. Tese de Doutorado. Programa integrado de pós-graduação em biologia tropical e recursos naturais. INPA/UFAM, 2008.
- QUEIROZ, W. T. **Análise multivariada em inventário florestal contínuo**. Belo Horizonte: Editora Poisson, 2021. 271 p. (<http://dx.doi.org/10.36229/978-65-5866-134-4>).
- SATTERTHWAITE, F.E. An approximate distribution of estimates of variance components. **Biometric Bulletin**, London, v. 2, p.110-114, 1946.
- SILVA, J. A. A.; SILVA, I, P. **Estatística experimental aplicada à ciência florestal**. Curso de engenharia Florestal. Recife: UFRPE, 1982. 300 p.
- STEEL, R. G. D.; TORRIE, J. H. **Principles and procedures of statistics: with special reference to the biological sciences**. New York: Mc Graw-Hill Book Company, Inc., 1960. 481 p.
- ZIMMERMANN, F. J. P. **Análise de um experimento de adubação e calagem em milho, executado num delineamento em faixas durante oito anos no mesmo local**. Dissertação de mestrado. Pós-graduação em estatística e experimentação agronômica. ESAQ-USP. Piracicaba, 1973. 93 p.

CAPÍTULO 9

DELINEAMENTOS EM BLOCOS INCOMPLETOS EQUILIBRADOS

Este capítulo apresenta definições, conceitos e aplicações do delineamento em blocos incompletos equilibrados (balanceados) e, ainda, a distribuição dos blocos que podem ser reunidos em cinco tipos distintos (I, II, III, IV e V). Apresenta também dois métodos de análise estatística: o primeiro chamado de Intrablocos e o segundo denominado Recuperação da Informação Interblocos. O capítulo, além disso, inclui exemplos de aplicação desses métodos.

9.1 INTRODUÇÃO

Os delineamentos em blocos incompletos equilibrados ou balanceados apresentam o mesmo número de unidades experimentais para todos os tratamentos, mas os blocos não incluem todos os tratamentos devido a algumas situações, pois, em algumas pesquisas florestais, as condições de campo muitas vezes restringem o uso de blocos completos.

Na experimentação florestal, podem ocorrer algumas situações que impossibilitam o uso do delineamento com a casualização dos tratamentos em blocos completos, tais como:

- a) o número de tratamentos no experimento é muito grande. Situação comum em pesquisas de melhoramento genético de espécies florestais, o que inviabiliza a formação de blocos homogêneos;
- b) o material experimental é muito heterogêneo;
- c) as condições do terreno apresentam limitações na instalação de determinados tamanhos de blocos.

Quando essas situações são verificadas, é indicado o uso do delineamento de blocos incompletos, cuja aplicação diminui a variância do erro experimental e propicia comparações mais precisas entre os tratamentos. Os delineamentos em blocos incompletos foram introduzidos por Yates (1936).

Pimentel Gomes e Garcia (2002) recomendam utilizar, na ocorrência de alguma limitação, os blocos incompletos equilibrados (*balanced incomplete blocks*), os reticulados quadrados (*square lattices* ou *two-dimensional lattices*) e os ensaios casualizados em blocos com alguns tratamentos comuns.

Neste livro, em especial, será dada ênfase ao delineamento em blocos incompletos balanceados (BIB), em que os I tratamentos ocorrem no mesmo bloco, com cada um dos demais, o mesmo número de vezes (λ). Nesse caso, o balanceamento é importante para estabelecer a mesma precisão na estimação dos efeitos dos tratamentos. Portanto, deve-se procurar manter o balanceamento, caso haja o interesse em efetuar todas as comparações entre tratamentos.

Basicamente, tomando-se $b = C_k^I$ blocos ($k < I$), pode-se planejar um BIB. Existem três condições necessárias para construir um experimento em blocos incompletos balanceados:

- a) $IJ = bk = n$;
- b) $J(k - 1) = \lambda(I - 1)$;
- c) $b \geq I$.

I : número de tratamentos;

b : número de blocos;

k : número de parcelas dentro dos blocos;

n : número total de parcelas experimentais;

J : número de vezes que cada tratamento está repetido.

Resultando em: $\lambda = \frac{J(k-1)}{I-1}$ ou $\lambda = \frac{bk(k-1)}{I(I-1)}$ ($\lambda < k < b$)

Tal que λ é o número de vezes em que cada par de tratamentos aparece no mesmo bloco.

Considerar o experimento com sete tratamentos, A, B, C, D, E, F, G ($I=7$), dispostos em sete blocos ($b=7$), com três parcelas ($k=3$) cada, totalizando vinte e uma unidades experimentais ou parcelas ($n=21$), em que cada tratamento está repetido três vezes ($J=3$).

Blocos	Trat.			Blocos	Trat.			Blocos	Trat.			Blocos	Trat.		
1	A	B	D	3	C	D	F	5	A	E	F	7	A	C	G
2	B	C	E	4	D	E	G	6	B	F	G	--	--	--	--

$$n = bk = JI = 7 \times 3 = 3 \times 7 = 21 \quad \lambda = \frac{J(k-1)}{I-1} = \frac{bk(k-1)}{I(I-1)}$$

$$\lambda = \frac{J(k-1)}{I-1} = \frac{3(3-1)}{7-1} = 1 \quad \lambda = \frac{bk(k-1)}{I(I-1)} = \frac{7 \times 3(3-1)}{7(7-1)} = 1$$

O fato de $\lambda = 1$ indica que qualquer tratamento aparece apenas uma vez combinado com os demais. Nesse exemplo, o par AB ocorre apenas uma vez (bloco B1), o par BC também apenas uma vez (bloco B2), assim acontecendo com os outros pares.

Considerar o planejamento experimental citado em Pimentel Gomes (2000), no qual há seis tratamentos ($I=6$), dispostos em dez blocos ($b=10$) com três parcelas ($k=3$), totalizando trinta unidades experimentais ($n=30$), em que cada tratamento está repetido cinco vezes ($J=5$).

	Trat.				Trat.				Trat.				Trat.				Trat.		
B1	A	B	E	B3	A	C	D	B5	A	D	E	B7	B	C	E	B9	C	E	F
B2	A	B	F	B4	A	C	F	B6	B	C	D	B8	B	D	F	B10	D	E	F

Em que:

$$n = bk = JI = 10 \times 3 = 5 \times 6 = 30$$

$$\lambda = \frac{J(k-1)}{I-1} = \frac{bk(k-1)}{I(I-1)}$$

$$\lambda = \frac{J(k-1)}{I-1} = \frac{5(3-1)}{6-1} = 2 \quad \lambda = \frac{bk(k-1)}{I(I-1)} = \frac{10 \times 3(3-1)}{6(6-1)} = 2$$

$$kJ = J + \lambda(I-1) = 5 + 2 \times (6-1) = 15$$

Como $\lambda = 2$, é possível indicar que qualquer tratamento aparece duas vezes combinado com os demais. Nesse exemplo, o par AB ocorre apenas nos blocos B1 e B2, o par BD aparece somente nos blocos B6 e B8, assim como os demais pares.

Os delineamentos em blocos incompletos são classificados em balanceados e parcialmente balanceados. Nos balanceados, cada par de tratamentos ocorre em um mesmo bloco o mesmo número de vezes como verificado nos exemplos anteriormente. Por outro lado, nos experimentos parcialmente balanceados, nem todos os pares aparecem o mesmo número de vezes juntos em um mesmo bloco. A condição de balanceamento é dada por:

$$\lambda(I-1) = J(k-1)$$

9.2. BLOCOS INCOMPLETOS BALANCEADOS (BIB)

Como já definido anteriormente:

I : número de tratamentos;

k : número de tratamentos por bloco;

b : número de blocos;

J : número de repetições dos tratamentos;

Cochran e Cox (1957) citam que os experimentos em blocos incompletos equilibrados podem ser reunidos em cinco tipos distintos (I, II, III, IV e V).

Tipo I: quando os blocos podem ser agrupados em repetições no experimento;

Tipo II: quando os blocos podem ser distribuídos em grupos de repetições;

Tipo III: quando não é possível agrupar os blocos em repetições ou grupos de repetições;

Tipo IV: é um caso particular do tipo III, em que o número de tratamentos é igual ao número de blocos ($I = b$);

Tipo V: quando aplicável a pequenos experimentos. Esse tipo sempre apresenta uma estrutura idêntica aos tipos I, II e III.

A seguir uma breve apresentação sobre os cinco tipos de experimentos em blocos incompletos:

a) **Tipo I:** quando os blocos podem ser agrupados em repetições (rep) no experimento.

$$x_{ij} = m + t_i + rep_j + (b/rep)_{ij} + e_{ij}$$

Definindo $b_j = rep_j + (b/rep)_{ij}$, então, $x_{ij} = m + t_i + b_j + e_{ij}$, que é o modelo usado em experimentos casualizados em blocos.

Considerar o plano experimental apresentado em Pimentel Gomes (2000), com oito tratamentos ($I = 8$) dispostos em vinte e oito blocos ($b = 28$) com duas parcelas ($k = 2$), totalizando cinquenta e seis parcelas ou unidades experimentais ($n = 56$), em que cada tratamento está repetido sete vezes ($J = 7$). Para facilitar a abordagem do tema, os exemplos que seguem serão chamados de planos:

Plano 9.1. BIB (Tipo I) com blocos agrupados em sete repetições.

Rep1		Rep2		Rep3		Rep4		Rep5		Rep6		Rep7	
(T1)	(T2)	(T1)	(T3)	(T1)	(T4)	(T1)	(T5)	(T1)	(T6)	(T1)	(T7)	(T1)	(T8)
(T3)	(T4)	(T2)	(T8)	(T2)	(T7)	(T2)	(T3)	(T2)	(T4)	(T2)	(T6)	(T2)	(T5)
(T5)	(T6)	(T4)	(T5)	(T3)	(T6)	(T4)	(T7)	(T3)	(T8)	(T3)	(T5)	(T3)	(T7)
(T7)	(T8)	(T6)	(T7)	(T5)	(T8)	(T6)	(T8)	(T5)	(T7)	(T4)	(T8)	(T4)	(T6)

$$\lambda = \frac{J(k-1)}{I-1} = \frac{7(2-1)}{8-1} = 1$$

Como $\lambda = 1$, depreende-se que qualquer tratamento aparece apenas uma vez combinado com os demais, resultando no seguinte quadro de análise de variância (ANOVA).

Tabela 9.2.1. ANOVA BIB (plano 9.1) com blocos ($b = 28$) agrupados em sete repetições.

F.V.	G.L.
Repetições (Rep.)	$J - 1 = 6$
Blocos d. Repetições	$b - J = 21$
Tratamentos (Aj.)	$I - 1 = 7$
Resíduo	$b(k-1) - (I-1) = 21$
Total	$(bk-1) = (IJ-1) = 55$

Considerar o planejamento experimental (plano 9.2), citado em Trevisol e Cosme (2013), em que há seis tratamentos ($I = 6$) dispostos em quinze blocos ($b = 15$) com duas parcelas cada ($k = 2$), totalizando trinta unidades experimentais ($n = 30$), em que cada tratamento está repetido cinco vezes ($J = 5$):

Plano 9.2. Plano experimental de BIB com blocos agrupados em cinco repetições.

Blocos	Rep1		Blocos	Rep2		Blocos	Rep3	
1	(T1)	(T2)	4	(T1)	(T3)	7	(T1)	(T4)
2	(T3)	(T4)	5	(T2)	(T5)	8	(T2)	(T6)
3	(T5)	(T6)	6	(T4)	(T6)	9	(T3)	(T5)
Blocos	Rep4		Blocos	Rep5				
10	(T1)	(T5)	13	(T1)	(T6)			
11	(T2)	(T4)	14	(T2)	(T3)			
12	(T3)	(T6)	15	(T4)	(T5)			

$$\lambda = \frac{J(k-1)}{I-1} = \frac{5(2-1)}{6-1} = 1$$

Tabela 9.2.2. ANOVA BIB (plano 9.2) com blocos agrupados em cinco repetições.

F.V.	G.L.
Repetições (R)	$J - 1 = 4$
Blocos d. Repetições	$b - J = 10$
Tratamentos (Aj.)	$I - 1 = 5$
Resíduo	$b(k-1) - (I-1) = 10$
Total	$(bk-1) = (IJ-1) = 29$

b) **Tipo II:** quando os blocos podem ser distribuídos em grupos de repetições.

Considerar o planejamento experimental (plano 9.3) apresentado em Pimentel Gomes (2000), contendo sete tratamentos ($I = 7$) dispostos em vinte e um blocos ($b = 21$) com duas parcelas cada ($k = 2$), distribuídos em três grupos de repetições ($g = 3$), totalizando quarenta e duas parcelas ($n = 42$), em que cada tratamento está repetido seis vezes ($J = 6$).

Plano 9.3. BIB com blocos distribuídos em grupos de repetições.

Blocos	Grupo1		Blocos	Grupo2		Blocos	Grupo3	
1	T1	T2	8	T1	T3	15	T1	T4
2	T2	T3	9	T3	T5	16	T4	T7
3	T3	T4	10	T5	T7	17	T7	T3
4	T4	T5	11	T7	T2	18	T3	T6
5	T5	T6	12	T2	T4	19	T6	T2
6	T6	T7	13	T4	T6	20	T2	T5
7	T7	T1	14	T6	T1	21	T5	T1

$$\lambda = \frac{J(k-1)}{I-1} = \frac{6(2-1)}{7-1} = 1$$

Como $\lambda = 1$, então qualquer tratamento aparece apenas uma vez combinado com os demais, resultando no seguinte quadro de análise de variância:

Tabela 9.2.3. ANOVA BIB (plano 9.3) com blocos distribuídos em grupos de repetições.

F.V.	G.L.
Grupos	$g - 1 = 2$
Blocos d. Grupos	$b - g = 18$
Tratamentos (Aj.)	$I - 1 = 6$
Resíduo	$b(k-1) - (I-1) = 15$
Total	$(bk-1) = (IJ-1) = 41$

$$x_{ij} = m + t_i + G_j + (b/G)_{ij} + e_{ij}$$

c) **Tipo III**: quando não é possível agrupar os blocos em repetições ou grupos de repetições

Considerar o planejamento experimental (plano 9.4) citado em Campos e Piedade (1992), em que seis tratamentos ($I = 6$) estão dispostos em dez blocos ($b = 10$) com três parcelas cada ($k = 3$), totalizando trinta unidades experimentais ($n = 30$), nas quais cada tratamento está repetido cinco vezes ($J = 5$):

Plano 9.4. BIB com blocos não distribuídos em repetições ou grupos de repetições.

Blocos	Tratamentos			Blocos	Tratamentos		
1	T1	T2	T5	6	T2	T3	T4
2	T1	T2	T6	7	T2	T3	T5
3	T1	T3	T4	8	T2	T4	T6
4	T1	T3	T6	9	T3	T5	T6
5	T1	T4	T5	10	T4	T5	T6

$$\lambda = \frac{J(k-1)}{I-1} = \frac{5(3-1)}{6-1} = 2$$

No plano 9.4, qualquer tratamento aparece duas vezes ($\lambda = 2$), no mesmo bloco com cada um dos demais. Por exemplo, os tratamentos T2 e T3 ocorrem duas vezes conjuntamente nos blocos B_6 e B_7 , os tratamentos T1 e T4 aparecem duas vezes nos blocos B_3 e B_5 . Por outro lado, cada tratamento aparece em cinco blocos, isto é, o número de repetições é igual a cinco ($J = 5$), resultando no seguinte quadro de análise de variância:

$$x_{ij} = m + t_i + b_j + e_{ij}$$

Tabela 9.2.4. ANOVA BIB (plano 9.4) com blocos não distribuídos em repetições ou em grupos de repetições.

F.V.	G.L.
Blocos	$b - 1 = 9$
Tratamentos (Aj.)	$I - 1 = 5$
Resíduo	$b(k - 1) - (I - 1) = 15$
Total	$(bk - 1) = (IJ - 1) = 29$

Considerar o planejamento experimental, com delineamento em blocos incompletos equilibrados do tipo III (plano 9.5), apresentado em Pimentel Gomes (2000), o qual apresenta cinco tratamentos ($I = 5$), dispostos em dez blocos ($b = 10$) com três parcelas cada ($k = 3$), resultando em trinta parcelas ou unidades experimentais ($n = 30$), em que cada tratamento está repetido seis vezes ($J = 6$).

Plano 9.5. BIB com blocos não distribuídos em repetições ou grupos de repetições.

Blocos	Tratamentos			Blocos	Tratamentos		
1	(T1)	(T2)	(T3)	6	(T1)	(T4)	(T5)
2	(T1)	(T2)	(T4)	7	(T2)	(T3)	(T4)
3	(T1)	(T2)	(T5)	8	(T2)	(T3)	(T5)
4	(T1)	(T3)	(T4)	9	(T2)	(T4)	(T5)
5	(T1)	(T3)	(T5)	10	(T3)	(T4)	(T5)

$$x_{ij} = m + t_i + b_j + e_{ij} \quad \lambda = \frac{J(k-1)}{I-1} = \frac{6(3-1)}{5-1} = 3$$

Tabela 9.2.5. ANOVA BIB (plano 9.5) com blocos não distribuídos em repetições ou grupos de repetições.

F.V.	G.L.
Blocos	$b - 1 = 9$
Tratamentos (Aj.)	$I - 1 = 4$
Resíduo	$b(k - 1) - (I - 1) = 16$
Total	$(bk - 1) = (IJ - 1) = 29$

d) **Tipo IV:** é um caso particular do tipo III, em que o número de tratamentos é igual ao número de blocos.

Abaixo, o exemplo do plano 9.6 apresentado por Cochran e Cox (1957), sendo $I = b = 13$ e $k = J = 4$.

Plano 9.6. BIB Tipo III com número de tratamentos igual ao de blocos.

Blocos	Tratamentos				Blocos	Tratamentos			
1	T3	T6	T9	T11	8	T3	T5	T7	T13
2	T3	T4	T8	T12	9	T1	T2	T3	T10
3	T10	T11	T12	T13	10	T2	T4	T9	T13
4	T2	T5	T8	T11	11	T1	T4	T7	T11
5	T7	T8	T9	T10	12	T1	T6	T8	T13
6	T4	T5	T6	T10	13	T2	T6	T7	T12
7	T1	T5	T9	T12					

Tabela 9.2.6. ANOVA BIB: tipo III (plano 9.6): número de tratamentos igual ao de blocos.

F.V.	G.L.
Blocos (Não Aj.)	$b - 1 = 12$
Tratamentos (Aj.)	$I - 1 = 12$
Resíduo	$b(k - 1) - (I - 1) = 27$
Total	$(bk - 1) = (IJ - 1) = 51$

$$\lambda = \frac{J(k-1)}{I-1} = \frac{4(4-1)}{13-1} = 1$$

e) **Tipo V**: é aplicável em pequenos experimentos e sempre apresenta um plano experimental correspondente a um dos tipos I, II ou III.

É o caso do exemplo (plano 9.7) mostrado por Campos e Piedade (1992), em que $b=6$, $I=4$, $k=2$ e $J=3$.

Plano 9.7. BIB do tipo V com planejamento do tipo I.

Blocos	Rep1		Blocos	Rep2		Blocos	Rep3	
1	T1	T2	3	T1	T3	5	T1	T4
2	T3	T4	4	T2	T4	6	T2	T3

$$\lambda = \frac{J(k-1)}{I-1} = \frac{3(2-1)}{4-1} = 1$$

Tabela 9.2.7. ANOVA BIB Tipo V (plano 9.7) com blocos agrupados em três repetições de acordo com o Tipo I.

F.V.	G.L.
Repetições	$J - 1 = 2$
Blocos d. repetições (Não Aj.)	$b - J = 3$
Tratamentos (Aj.)	$I - 1 = 3$
Resíduo	$b(k-1) - (I-1) = 3$
Total	$(bk-1) = (IJ-1) = 11$

O uso de delineamentos em blocos balanceados propicia a mesma precisão para todas as comparações entre tratamentos. Quando os recursos são limitados e o número de repetições necessárias é muito alto, e não existindo unidades experimentais suficientes para construir um delineamento em blocos incompletos balanceados, é possível estruturar delineamentos parcialmente balanceados (BIPB) com um menor número de repetições. Os desenhos de blocos incompletos parcialmente balanceados, ao contrário dos balanceados, possuem alguns pares de tratamentos, os quais aparecem em mais blocos que em outros pares. Consequentemente, algumas comparações de tratamentos são mais precisas que outras.

A seguir, o delineamento de blocos incompletos parcialmente balanceados (plano 9.8), citado em Trevisol e Cosme (2013), para seis tratamentos em três blocos com quatro parcelas:

Plano 9.8. Exemplo de um delineamento BIPB.

Bloco1	T1	T4	T2	T5
Bloco2	T2	T5	T3	T6
Bloco3	T3	T6	T1	T4

Nota-se que os pares de tratamentos (T1, T4), (T2, T5) e (T3, T6) aparecem em dois blocos ($\lambda = 2$), enquanto os demais pares aparecem em apenas um bloco ($\lambda = 1$). Os pares de tratamentos que aparecem juntos em dois blocos são comparados com precisão um pouco maior do que os que estão em apenas um. Por outro lado, as diferenças de precisão, normalmente, não são tão relevantes para não adotar o uso do delineamento parcialmente balanceado.

Não obstante, em alguns casos, na experimentação florestal, é impossível construir delineamentos balanceados que requerem blocos incompletos, pois, em diversas situações, o número de repetições necessárias dificulta o estabelecimento das unidades experimentais suficientes para implantar o experimento. Portanto, normalmente são construídos delineamentos parcialmente balanceados que demandam menos repetições.

Sempre existe, teoricamente, para I tratamentos e k parcelas por bloco, um plano experimental em blocos incompletos balanceado com $b = C_k^I$ números de blocos. Por exemplo, para um delineamento em blocos incompletos balanceados (BIB), apresentando $I = 6$ tratamentos com $k = 4$ parcelas por bloco, o número de blocos será $C_k^I = C_4^6 = 15$, requer o estabelecimento de sessenta parcelas ou unidades experimentais. Entretanto, como $b \times k = 15 \times 4 = 60$ e $b \times k = IJ = 60$, o número de repetições será igual a dez ($J = 10$) e, consequentemente, $\lambda = 6$.

$$\lambda = \frac{J(k-1)}{I-1} = \frac{10(4-1)}{6-1} = 6$$

No caso do delineamento parcialmente balanceado (BIPB plano 9.8) com $I = 6$ e $k = 4$, no qual ocorreu uma grande redução no número de blocos e no número de repetições, nota-se que são necessárias apenas doze unidades experimentais.

Bloco1	T1	T4	T2	T5
Bloco2	T2	T5	T3	T6
Bloco3	T3	T6	T1	T4

9.3. ANÁLISE ESTATÍSTICA DE BLOCOS INCOMPLETOS EQUILIBRADOS (BALANCEADOS)

A análise dos experimentos em blocos incompletos é um pouco mais difícil do que a dos experimentos em blocos completos. Contudo, a análise mais complicada dos BIBs, como também a perda de graus de liberdade do resíduo, muitas vezes, é compensada por uma redução na soma de quadrados do resíduo, influenciando seu quadrado médio do erro, resultando na obtenção de experimentos com comparações de médias mais precisas.

Existem dois tipos de análise de experimentos em blocos incompletos: análise intrablocos e a análise com recuperação da informação interblocos, aplicáveis independentemente do tipo de experimento.

a) Na análise intrablocos, introduzida por Yates (1936), as comparações são feitas entre parcelas do mesmo bloco, isto é, apenas essas são utilizadas nas estimativas de efeito de tratamentos, pois o efeito de cada bloco é considerado fixo. Nesse método qualquer interferência de efeito de bloco é descartada. Tal procedimento é utilizado para qualquer experimento em blocos incompletos e se baseia em métodos estatísticos exatos.

Considerar o modelo estatístico usado na análise intrablocos para representar x_{ij} , que corresponde ao valor da variável resposta (x), a qual recebeu o i -ésimo tratamento e pertence ao j -ésimo bloco:

$$x_{ij} = m + t_i + b_j + e_{ij}$$

m : média geral;

t_i : efeito do i -ésimo tratamento;

b_j : efeito do j -ésimo bloco;

e_{ij} : erro experimental, tal que $e_{ij} \sim N(0, \sigma^2)$.

b) Na análise com recuperação da informação interblocos apresentada por Yates (1940), as comparações entre blocos também são utilizadas na estimação dos efeitos dos tratamentos. De acordo com Trevisol e Cosme (2013), além das comparações feitas dentro de cada bloco, as comparações entre blocos são aproveitadas para a estimação dos efeitos dos tratamentos — ou seja, a partir de duas estimativas diferentes dos efeitos de tratamentos — se obtiverem uma terceira estimativa. Esse método, de acordo com Pimentel Gomes (2000), aprofunda melhor a análise dos dados, mas utiliza métodos estatísticos apenas aproximados. É recomendado para a análise de experimentos com um número de graus de liberdade relativamente grande para os blocos e para o resíduo.

$$x_{ij} = m + t_i + b_j + e_{ij}$$

m : média geral;

t_i : efeito do i -ésimo tratamento;

b_j : efeito aleatório do j -ésimo bloco, tal que $b_j \sim N(0, \sigma_b^2)$;

e_{ij} : erro experimental, tal que $e_{ij} \sim N(0, \sigma^2)$.

É interessante obter o valor da eficiência (E) dos experimentos em blocos incompletos equilibrados ou balanceados (BIB) em relação aos blocos completos, a qual é definida pela fórmula:

$$E = \frac{\lambda I}{k J} \rightarrow 0 < E < 1$$

9.3.1 Análise intrablocos em blocos incompletos equilibrados (balanceados)

Modelo matemático:

$$x_{ij} = m + t_i + b_j + e_{ij}$$

m : média geral;

t_i : efeito do i -ésimo tratamento;

b_j : efeito do j -ésimo bloco considerado fixo;

e_{ij} : erro experimental, tal que $e_{ij} \sim N(0, \sigma^2)$.

O efeito de blocos é considerado fixo na análise intrablocos dos experimentos em blocos incompletos. A fim de obter a soma de quadrados para os tratamentos ajustados para os efeitos de blocos, segundo Pimentel Gomes (2000), os seguintes cálculos devem ser efetuados, de acordo com os seguintes passos:

- Obter os totais por tratamentos T_i ($i = 1, \dots, I$);
- Calcular as correções (A_i) para os efeitos de blocos, em que A_i é a soma dos totais dos blocos quando aparece o correspondente i -ésimo tratamento;
- Obter os valores $Q_i = kT_i - A_i$;
- Estimar os valores dos efeitos ajustados de tratamento ($\hat{t}_i$), em que:

$$\hat{t}_i = \frac{Q_i}{\lambda I}, \text{ tal que: } \hat{m}_i = \hat{m} + \hat{t}_i$$

- Calcular a soma de quadrados para tratamentos ajustados pela fórmula:

$$SQ_{TRAT(A_j)} = \frac{\sum_{i=1}^I Q_i^2}{\lambda k I}$$

- Fazer comparações múltiplas.

Nos ensaios em blocos completos, a fórmula para obter a estimativa da variância da diferença entre as estimativas das médias entre dois tratamentos, sendo $\hat{\sigma}^2$ a estimativa da variância residual, é:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2(QM_{RES})}{J} = \frac{2\hat{\sigma}^2}{J}$$

Nos ensaios em blocos incompletos equilibrados, a fórmula para obter a estimativa da variância da diferença entre as estimativas das médias entre dois tratamentos é dada por:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{1}{E} \frac{2\hat{\sigma}^2}{J} \quad E = \frac{\lambda I}{k J} \quad \hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2k}{\lambda I} \hat{\sigma}^2$$

Desse modo, maior será a precisão das estimativas quanto mais próximo de 1 (um) for o valor da

eficiência (E).

Para comparação de médias na aplicação do teste de Tukey:

$$\Delta = q \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)}$$

No caso de um contraste da forma $\hat{Y} = c_1 \hat{m}_1 + c_2 \hat{m}_2 + \dots + c_I \hat{m}_I$, tem-se:

$$\hat{V}(\hat{Y}) = \frac{k}{\lambda I} \left(\sum_{i=1}^I c_i^2 \right) \hat{\sigma}^2$$

Alguns exemplos de aplicação de análise intrablocos para os tipos I, II e III de blocos incompletos equilibrados são apresentados a seguir:

Exercício 9.3.1.1. Tipo I: De acordo com o delineamento apresentado no plano 9.2, considerar o ensaio com seis métodos de manejo de uma espécie florestal ($I=6$), (M1, M2, M3, M4, M5, M6). O experimento utilizado foi de blocos incompletos equilibrados do tipo I, considerando quinze blocos ($b=15$) com duas parcelas ($k=2$) de meio hectare em cada, agrupados em cinco repetições. A variável resposta (dados simulados) foi o volume de madeira ao fim da rotação de dez anos. Observa-se, na configuração do ensaio (plano 9.2), que qualquer tratamento aparece uma única vez ($\lambda=1$), no mesmo bloco, com cada um dos demais. O número de repetições é igual a cinco ($J=5$). Efetuar a análise pelo método intrablocos.

Tabela 9.3.1.1. Plano experimental de BIB com blocos agrupados em cinco repetições.

Rep1		Total	Rep2		Total	Rep3		Total
36,6 (M1)	32,9 (M2)	69,5	43,9 (M1)	32,8 (M3)	76,7	42,0 (M1)	31,1 (M4)	73,1
27,4 (M3)	29,7 (M4)	57,1	45,7 (M2)	29,2 (M5)	74,9	47,5 (M2)	31,7 (M6)	79,2
25,6 (M5)	27,4 (M6)	53,0	23,7 (M4)	21,9 (M6)	45,6	27,4 (M3)	23,7 (M5)	51,1
179,6			197,2			203,4		
Rep4		Total	Rep5		Total			
38,4 (M1)	23,7 (M5)	62,1	51,2 (M1)	25,6 (M6)	76,8			
42,0 (M2)	18,3 (M4)	60,3	49,4 (M2)	33,9 (M3)	83,3			
29,42 (M3)	23,7 (M6)	53,1	32,5 (M4)	29,5 (M5)	62,0			
175,5			222,1					

a) Eficiência (E) e λ :

$$\lambda = \frac{J(k-1)}{I-1} = \frac{5(2-1)}{6-1} = 1 \quad E = \frac{\lambda I}{kJ} = \frac{1 \times 6}{2 \times 5} = 0,6 \rightarrow 60,0\%$$

b) Somas de quadrados:

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 34.105,28 - \frac{977,8^2}{6 \times 5} = 2.235,5187$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{k} (B_1^2 + B_2^2 + \dots + B_b^2) - C$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{2} (69,5^2 + 57,1^2 + \dots + 83,3^2 + 62,0^2) - C$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{2} (65.692,62) - \frac{977,8^2}{30} = 976,5487$$

$$SQ_{REP} = \frac{1}{6} (179,6^2 + 197,2^2 + 203,4^2 + 175,5^2 + 222,1^2) - \frac{977,8^2}{30}$$

$$SQ_{REP} = \frac{1}{6} (192.644,22) - \frac{977,8^2}{30} = 237,6087$$

$$SQ_{BLOCOS D.REP(N\tilde{a}o Aj.)} = SQ_{BLOCOS} - SQ_{REP} = 976,5487 - 237,6087 = 738,9400$$

c) Soma de quadrados para tratamentos ajustados para os efeitos de blocos. A fim de obter $SQ_{TRAT(Aj.)}$, é necessário calcular:

c1) totais por tratamentos T_i ($i=1, \dots, I$)

T1	212,1	T3	150,9	T5	131,7
T2	217,5	T4	135,3	T6	130,3

c2) soma total dos blocos em que aparece o tratamento i correspondente (A_i):

A1	69,5+76,7+73,1+62,1+76,8	358,2	A4	57,1+45,6+73,1+60,3+62,0	298,1
A2	69,5+74,9+79,2+60,3+83,3	367,2	A5	53,0+74,9+51,1+62,1+62,0	303,1
A3	57,1+76,7+51,1+53,1+83,3	321,3	A6	53,0+45,6+79,2+53,1+76,8	307,7

c3) calcular $Q_i = kT_i - A_i$

Então, a soma de quadrados de tratamentos ajustados é obtida a partir dos dados disponibilizados na tabela 9.3.1.2:

$$SQ_{TRAT(Aj.)} = \frac{1}{\lambda k I} \sum_{i=1}^I Q_i^2 = \frac{1}{1 \times 2 \times 6} (13.883,84) = 1.156,9867$$

d) Realizações dos testes de comparações múltiplas, sendo necessário estimar os efeitos ajustados dos tratamentos ($\hat{t}_i$) e, assim, obter as suas médias ajustadas ($\hat{m}_i$). Os valores de $\hat{t}_i$ e $\hat{m}_i$ estão apresentados na tabela 9.3.1.2.

$$\hat{t}_i = \frac{1}{\lambda J} Q_i = \frac{Q_i}{6} \quad \hat{m}_i = \hat{m} + \hat{t}_i \quad \hat{m} = \frac{\sum_{i=1}^I \sum_{j=1}^J x_{ij}}{IJ} = \frac{977,8}{30} = 32,5933$$

Tabela 9.3.1.2. Resultados dos cálculos para a análise intrablocos.

Tratamentos	T_i	A_i	$Q_i = kT_i - A_i$	$\hat{t}_i = \frac{1}{\lambda J} Q_i$	$\hat{m}_i = \hat{m} + \hat{t}_i$
M1	212,10	358,2	66,0	11,0000	43,5933
M2	217,50	367,2	67,8	11,3000	43,8933
M3	150,90	321,3	-19,5	-3,2500	29,3433
M4	135,30	298,1	-27,5	-4,5833	28,0100
M5	131,70	303,1	-39,7	-6,6167	25,9766
M6	130,30	307,7	-47,1	-7,8500	24,7433
			0	0	

e) Análise de variância pelo método de análise intrablocos.

Tabela 9.3.1.3. ANOVA intrablocos.

F.V.	G.L.	S.Q.	Q.M.	F
Repetição (Rep.)	4	237,6087	59,4022	5,82 *
Blocos d. Rep. (Não Aj.)	10	738,9400	73,8940	7,25 **
Trat. (Aj.)	5	1.156,9867	231,3973	22,69 **
Resíduo	10	101,9833	10,1983	
Total	29	2.235,5187		

$$F_{5;10;0,01} = 5,64 \quad F_{10;10;0,01} = 4,85 \quad F_{4;10;0,01} = 5,99 \quad F_{4;10;0,05} = 3,48$$

Tabela 9.3.1.4. ANOVA intrablocos no SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Repetições (Rep.)	4	237,608667	59,402167	5,82	0,0110
Blocos d. Rep. (Não Aj.)	10	738,940000	73,894000	7,25	0,0022
Trat. (Aj.)	5	1156,986667	231,397333	22,69	< 0,0001
Resíduo	10	101,983333	10,198333		
Total	29	2235,518667			

O programa **Prog26**, escrito no sistema de análise estatística SAS, em anexo no apêndice B, efetua a análise intrablocos do experimento de blocos incompletos equilibrados do tipo I.

f) Teste de comparação múltipla de Tukey

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2k}{\lambda I} QM_{\text{Res}} = \frac{2 \times 2}{1 \times 6} \times 10,1983 = 6,7989$$

$$\Delta = q_{(6; 10; 0,05)} \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}_i - \hat{m}_u)} \quad \Delta = 4,91 \sqrt{\left(\frac{1}{2}\right) 6,7956} = 9,05$$

Tratamentos	Médias Ajustadas
M2	43,89 a
M1	43,59 a
M3	29,34 b
M4	28,01 b
M5	25,98 b
M6	24,74 b

Médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem pelo teste de Tukey. Os métodos de manejo M2 e M1 apresentaram um comportamento diferenciado em relação aos demais.

Exercício 9.3.1.2. Tipo II. De acordo com o delineamento apresentado no plano 9.3, considerar um experimento com sete métodos de manejo florestal de uma espécie florestal ($I = 7$), (M1, M2, M3, M4, M5, M6, M7). O delineamento utilizado foi de blocos incompletos equilibrados do tipo II, considerando vinte e um blocos ($b = 21$), com duas parcelas ($k = 2$) de meio hectare em cada, distribuídos em três grupos de repetições. A variável resposta medida (dados simulados) foi o volume de madeira ao final da rotação de dez anos. Verifica-se, na configuração do experimento, que qualquer tratamento aparece uma única vez ($\lambda = 1$), no mesmo bloco, com cada um dos demais. O número de repetições é igual a seis ($J = 6$). Efetuar a análise intrablocos.

Tabela 9.3.1.5. Dados do delineamento em bloco incompleto equilibrados do tipo II.

GRUPO1			Totais	GRUPO2			Totais	GRUPO3			Totais
M1=81,5	M2=65,2	146,7		M1=88,5	M3=93,2	181,7		M1=69,9	M4=51,2	121,1	
M2=74,5	M3=86,2	160,7		M3=83,8	M5=55,9	139,7		M4=62,9	M7=79,2	142,1	
M3=81,5	M4=58,2	139,7		M5=46,6	M7=69,9	116,5		M7=74,5	M3=90,8	165,3	
M4=65,2	M5=55,9	121,1		M7=65,2	M2=58,2	123,4		M3=76,8	M6=55,9	132,7	
M5=62,9	M6=74,5	137,4		M2=60,5	M4=65,2	125,7		M6=65,2	M2=79,2	144,4	
M6=55,9	M7=60,5	116,4		M4=53,5	M6=55,9	109,4		M2=67,5	M5=53,5	121,0	
M7=72,2	M1=62,9	135,1		M6=65,2	M1=76,8	142,0		M5=46,6	M1=76,8	123,4	
		957,1				938,4				950,0	

a) Eficiência (E) e o valor de λ :

$$E = \frac{\lambda I}{k J} = \frac{1 \times 7}{2 \times 6} = 0,5833 \rightarrow 58,33\% \quad \lambda = \frac{J(k-1)}{I-1} = \frac{6(2-1)}{7-1} = 1$$

b) Somas de quadrados:

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 198715,93 - \frac{2.845,5^2}{7 \times 6} = 5.933,3050$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{k} (B_1^2 + B_2^2 + \dots + B_b^2) - C$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{2} (146,7^2 + 160,7^2 + \dots + 121,0^2 + 123,4^2) - \frac{2.845,5^2}{42}$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{2} (391.980,97) - \frac{2.845,5^2}{42} = 3.207,86$$

$$SQ_{GRUPOS} = \frac{1}{14} (957,1^2 + 938,4^2 + 950,0^2) - \frac{2.845,5^2}{42}$$

$$SQ_{GRUPOS} = \frac{1}{14} (2.699.134,97) - \frac{2.845,5^2}{42} = 12,73$$

$$SQ_{BLOCOS D. GRUPOS} = SQ_{BLOCOS} - SQ_{GRUPOS} = 3.207,86 - 12,73 = 3.195,13$$

c) Cálculo da soma de quadrados para tratamentos ajustados para os efeitos de blocos. A fim de obter $SQ_{TRAT(Aj.)}$, as seguintes informações são necessárias:

c1) os totais por tratamentos T_i ($i=1, \dots, I$)

T1	456,4	T3	512,3	T5	321,4	T7	421,5
T2	405,1	T4	356,2	T6	372,6		

c2) a soma total dos blocos em que aparece o tratamento i correspondente (A_i):

A1	146,7+181,7+121,1+135,1+142,0+123,4	850,0
A2	146,7+160,7+123,4+125,7+144,4+121,0	821,9
A3	181,7+160,7+139,7+139,7+165,3+132,7	919,8
A4	121,1+142,1+139,7+121,1+125,7+109,4	759,1
A5	139,7+116,5+121,1+137,4+121,0+123,4	759,1
A6	132,7+137,4+144,4+116,4+109,4+142,0	782,3
A7	142,1+116,5+165,3+123,4+116,4+135,1	798,8

c3) o cálculo de $Q_i = kT_i - A_i$

A soma de quadrados de tratamentos ajustados é obtida a partir dos cálculos disponibilizados na tabela 9.3.1.6.

$$SQ_{TRAT(Aj.)} = \frac{1}{\lambda k I} \sum_{i=1}^I Q_i^2 = \frac{1}{1 \times 2 \times 7} \times 34.100,4 = 2.435,7429$$

Tabela 9.3.1.6. Cálculos para a análise intrablocos.

	T_i	A_i	$Q_i = kT_i - A_i$	$\hat{t}_i = \frac{1}{\lambda I} Q_i$	$\hat{m}_i = \hat{m} + \hat{t}_i$
M1	456,4	850,0	62,8	8,9714	76,7214
M2	405,1	821,9	-11,7	-1,6714	66,0786
M3	512,3	919,8	104,8	14,9714	82,7214
M4	356,2	759,1	-46,7	-6,6714	61,0786
M5	321,4	759,1	-116,3	-16,6143	51,1357
M6	372,6	782,3	-37,1	-5,3000	62,4500
M7	421,5	798,8	44,2	6,3143	74,0643
Total			0	0	

d) Na realização dos testes de comparações múltiplas, é necessário estimar os efeitos ajustados dos tratamentos ($\hat{t}_i$) e obter as suas médias ajustadas ($\hat{m}_i$). Os valores de $\hat{t}_i$ e $\hat{m}_i$ estão apresentados na tabela 9.3.1.6.

$$\hat{t}_i = \frac{1}{\lambda I} Q_i$$

$$\hat{m} = \frac{\sum_{i=1}^I \sum_{j=1}^J x_{ij}}{IJ} = \frac{2.845,5}{42} = 59,2813$$

$$\hat{m}_i = \hat{m} + \hat{t}_i \quad \hat{m}_i = 59,2813 + \hat{t}_i$$

Tabela 9.3.1.7. ANOVA da análise intrablocos.

F.V.	G.L.	S.Q.	Q.M.	F
Grupos	2	12,7300	6,3650	0,33 <i>ns</i>
Blocos d. grupos (Não Aj.)	18	3.195,1300	177,5072	9,19**
Trat. (Aj.)	6	2.435,7429	405,9571	21,02**
Resíduo	15	289,7021	19,3135	
Total	41	5.933,3050		

$$F_{2;15;0,05} = 3,68 \quad F_{6;15;0,05} = 2,79 \quad F_{18;15;0,05} = 3,43$$

O programa **Prog27**, escrito no sistema de análise estatística SAS, em anexo no apêndice B, efetua a análise intrablocos para o ensaio em blocos incompletos equilibrados do tipo II.

Tabela 9.3.1.8. ANOVA da análise intrablocos pelo SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Grupos	2	12,730000	6,365000	0,33	0,7243
Blocos d. Grupos (Não Aj.)	18	3195,130000	177,507222	9,19	< 0,0001
Trat. (Aj.)	6	2435,742857	405,957143	21,02	< 0,0001
Resíduo	15	289,702143	19,313476		
Total	41	5933,305000			

e) O teste de comparação múltipla de Tukey

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2k}{\lambda I} QM_{\text{Res}} = \frac{2 \times 2}{1 \times 7} \times 19,3135 = 11,0363$$

$$\Delta = q_{(7; 15; 0,05)} \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}_i - \hat{m}_u)}$$

$$\Delta = 4,78 \sqrt{\left(\frac{1}{2}\right) 11,0363} = 11,2286$$

Tratamentos	Médias Ajustadas
M3	82,72 a
M1	76,72 ab
M7	74,06 ab
M2	66,08 bc
M6	62,45 c
M4	61,08 cd
M5	51,13 d

Médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem pelo teste de Tukey. Os métodos de manejo M3, M1 e M7 apresentaram um comportamento diferenciado em relação aos demais. O método M3 é o recomendável.

Exercício 9.3.1.3. Tipo III. De acordo com o delineamento mostrado no plano 9.6, considerar um experimento de competição de treze procedências de uma espécie florestal, $I = 13$, (P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13). O delineamento utilizado foi em blocos incompletos equilibrados do tipo III, considerando treze blocos ($b = 13$), com quatro parcelas ($k = 4$) de meio hectare em cada,

distribuídos em quatro repetições ($J = 4$). A variável resposta medida (dados simulados) foi o volume de madeira ao final da rotação de dez anos. Efetuar a análise pelo método intrablocos. Na tabela 9.3.1.9, estão os dados para análise intrablocos considerando um experimento do tipo III.

$$\lambda = \frac{J(k-1)}{I-1} = \frac{4(4-1)}{13-1} = 1 \quad E = \frac{\lambda I}{kJ} = \frac{1 \times 13}{4 \times 4} = 0,8125 \rightarrow 81,25\%$$

Tabela 9.3.1.9. Dados para análise de recuperação intrablocos experimento tipo III.

Blocos	Volume de madeira por procedência (m³ por 0,5 ha)								Total
1	P3	54,0	P6	42,4	P9	61,9	P11	42,5	200,8
2	P3	49,1	P4	42,2	P8	76,1	P12	48,4	215,8
3	P10	34,0	P11	31,2	P12	67,6	P13	66,8	199,6
4	P2	58,3	P5	57,6	P8	81,0	P11	27,1	224,0
5	P7	49,9	P8	70,1	P9	65,7	P10	69,1	254,8
6	P4	65,3	P5	69,1	P6	58,0	P10	70,0	262,4
7	P1	74,1	P5	66,4	P9	54,8	P12	65,1	260,4
8	P3	73,4	P5	69,1	P7	71,1	P13	88,8	302,4
9	P1	81,5	P2	70,2	P3	79,6	P10	66,8	298,1
10	P2	61,2	P4	65,5	P9	57,4	P13	85,3	269,4
11	P1	78,1	P4	66,4	P7	66,4	P11	50,6	261,5
12	P1	67,9	P6	71,9	P8	64,3	P13	97,7	301,8
13	P2	64,7	P6	67,2	P7	83,9	P12	57,0	272,8

a) Cálculo das somas de quadrados

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 222.795,66 - \frac{3.323,8^2}{13 \times 4} = 10.340,9208$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{4} (200,8^2 + 215,8^2 + \dots + 301,8^2 + 272,8^2) - \frac{3.323,8^2}{52} = 3.860,8758$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{4} (865.262,46) - \frac{3.323,8^2}{52} = 3.860,8758$$

$$SQ_{TRAT (N\tilde{a}o Aj.)} = \frac{1}{J} (T_1^2 + T_2^2 + \dots + T_I^2) - C$$

Os valores de T_i ($i=1, \dots, I$) estão mostrados na tabela 9.3.1.10.

$$SQ_{TRAT (N\tilde{a}o Aj.)} = \frac{1}{4} (301,6^2 + 254,4^2 + \dots + 238,1^2 + 338,6^2) - C$$

$$SQ_{TRAT(Não Aj.)} = \frac{1}{4}(872.586,1) - \frac{3.323,8^2}{52} = 5.691,7858$$

b) Obtenção da soma de quadrados para tratamentos ajustados para os efeitos de blocos $SQ_{TRAT(Aj.)}$, sendo necessário calcular:

b1) os totais por tratamentos T_i ($i=1, \dots, I$) estão mostrados na tabela 9.3.1.10.

b2) a soma total dos blocos em que aparece o tratamento i correspondente (A_i):

		A_i
P1	260,4+298,1+261,5+301,8	1.121,8
P2	224,0+269,4+272,8+298,1	1.064,3
P3	200,8+215,8+302,4+298,1	1.017,1
P4	262,4+215,8+269,4+261,5	1.009,1
P5	224,0+262,4+260,4+302,4	1.049,2
P6	200,8+301,8+272,8+262,4	1.037,8
P7	254,8+302,4+261,5+272,8	1.091,5
P8	254,8+215,8+224,0+301,8	996,4
P9	200,8+254,8+260,4+269,4	985,4
P10	199,6+254,8+262,4+298,1	1.014,9
P11	199,6+200,8+224,0+261,5	885,9
P12	199,6+215,8+260,4+272,8	948,6
P13	199,6+302,4+269,4+301,8	1.073,2

b3) obter $Q_i = kT_i - A_i$

Tabela 9.3.1.10. Resultados dos cálculos para a análise intrablocos.

	T_i	A_i	$Q_i = kT_i - A_i$	$\hat{t}_i$	$\hat{m}_i = \hat{m} + \hat{t}_i$
P1	301,6	1.121,8	84,6	6,5077	70,4269
P2	254,4	1.064,3	-46,7	-3,5923	60,3269
P3	256,1	1.017,1	7,3	0,5615	64,4807
P4	239,4	1.009,1	-51,5	-3,9615	59,9577
P5	262,2	1.049,2	-0,4	-0,0308	63,8884
P6	239,5	1.037,8	-79,8	-6,1385	57,7807
P7	271,3	1.091,5	-6,3	-0,4846	63,4346
P8	291,5	996,4	169,6	13,0462	76,9654
P9	239,8	985,4	-26,2	-2,0154	61,9038
P10	239,9	1.014,9	-55,3	-4,2538	59,6654
P11	151,4	885,9	-280,3	-21,5615	42,3577
P12	238,1	948,6	3,8	0,2923	64,2115
P13	338,6	1.073,2	281,2	21,6308	85,5500
			0	0	

A soma de quadrados de tratamentos ajustados é obtida a partir dos dados disponibilizados na tabela 9.3.1.10:

$$SQ_{TRAT(Aj)} = \frac{1}{\lambda KI} \sum_{i=1}^I Q_i^2 = \frac{1}{1 \times 4 \times 13} 208.616,14 = 4.011,8488$$

c) Realização dos testes de comparações múltiplas, sendo necessário estimar os efeitos ajustados dos tratamentos ($\hat{t}_i$) e, assim, obter as suas médias ajustadas ($\hat{m}_i$). Os valores de $\hat{t}_i$ e $\hat{m}_i$ estão apresentados na tabela 9.3.1.10.

$$\hat{t}_i = \frac{1}{\lambda I} Q_i = \frac{Q_i}{1 \times 13}$$

$$\hat{m} = \frac{\sum_{i=1}^I \sum_{j=1}^J x_{ij}}{IJ} = \frac{3.323,8}{52} = 63,9192$$

$$\hat{m}_i = \hat{m} + \hat{t}_i = 63,9192 + \hat{t}_i$$

Tabela 9.3.1.11. ANOVA intrablocos BIB tipo III.

F.V.	G.L.	S.Q.	Q.M.	F
Blocos (Não Aj.)	12	3.860,8758	321,7396	3,52**
Trat. (Aj)	12	4.011,8488	334,3207	3,66**
Resíduo	27	2.468,1962	91,4147	
Total	51	10.340,9208		

$$F_{12; 27; 0,05} = 2,10 \quad F_{12; 27; 0,01} = 2,93$$

Tabela 9.3.1.12. ANOVA intrablocos no SAS (BIB tipo III).

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Blocos (Não Aj.)	12	3860,875769	321,739647	3,52	0,0032
Trat.(Aj.)	12	4011,848846	334,320737	3,66	0,0025
Resíduo	27	2468,19615	91,41467		
Total	51	10340,92077			

O programa **Prog28**, construído no sistema de análise estatística SAS, em anexo no apêndice B, realiza a análise intrablocos para o experimento em blocos incompletos equilibrados do tipo III.

d) Teste de comparação múltipla de Tukey

$$\Delta = q_{(13; 27; 0,05)} \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}_i - \hat{m}_u)} \quad q_{(13; 27; 0,05)} \cong 5,13$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2k}{\lambda I} QM_{\text{Res}} = \frac{2 \times 4}{1 \times 13} \times 91,4147 = 56,2552$$

$$\Delta = 5,13 \sqrt{\left(\frac{1}{2}\right) 56,2552} = 27,21$$

Tratamentos	Médias ajustadas
P13	85,55 a
P8	76,96 ab
P1	70,43 ab
P3	64,48 abc
P12	64,21 abc
P5	63,89 abc
P7	63,43 abc
P2	60,33 abc
P9	61,90 abc
P4	59,96 abc
P10	59,66 abc
P6	57,78 bc
P11	42,36 c

Médias seguidas pela mesma letra não diferem entre si, pelo teste de Tukey, ao nível de 5% de significância. A procedência P13 é a recomendável.

9.3.2 Análise com recuperação da informação interblocos em blocos incompletos equilibrados (balanceados)

Os efeitos de blocos na análise intrablocos são considerados fixos e, conseqüentemente, os contrastes entre eles não são utilizados na estimação dos efeitos dos tratamentos. Entretanto, se os efeitos dos blocos forem considerados aleatórios, é possível combinar essa informação com as fornecidas, na análise intrablocos, pelos contrastes de parcelas dentro dos blocos e assim obter estimativas mais precisas dos efeitos de tratamentos.

Considerar o modelo matemático do experimento tipo III, isto é, em que os blocos não estão agrupados em repetições ou grupos de repetições:

$$x_{ij} = m + b_j + t_i + e_{ij}$$

m : média geral;

b_j : efeito aleatório do j -ésimo bloco com variância σ_b^2 ;

t_i : efeito fixo do i -ésimo tratamento;

e_{ij} : variável aleatória, com distribuição normal com variância σ^2 .

As estimativas de σ^2 e σ_b^2 são obtidas a partir do quadrado médio do resíduo ($V_r = \sigma^2$) e do quadrado médio para blocos (ajustados) (V_b). Para proceder à análise interblocos com recuperação da informação, Pimentel Gomes (2000) apresenta a seguinte sequência de cálculos:

a) Obter o valor: $a = \frac{\sigma^2}{\sigma^2 + k\sigma_b^2} \quad (0 < a < 1)$

No caso de $a = 1$, ocorre $\sigma_b^2 = 0$. Por outro lado, se σ_b^2 for muito grande em relação à σ^2 , então o valor de (a) será próximo de zero.

Sejam, abaixo, as expressões para o valor (a) , considerando os três tipos de blocos incompletos:

Tipo I: quando os blocos podem ser agrupados em repetições no experimento:

$$a = \frac{(J-1)\sigma^2}{JV_b - \sigma^2} \quad V_r = \sigma^2$$

Tipo II: quando os blocos podem ser distribuídos em grupos de repetições:

$$a = \frac{[I(J-1) - k(g-1)]\sigma^2}{k(b-g)V_b - (I-k)\sigma^2} \quad V_r = \sigma^2$$

Tipo III: quando não é possível agrupar os blocos em repetições ou grupos de repetições.

Pimentel Gomes (2000) mostra que, no caso dos experimentos tipo III, tem-se:

$$\sigma_b^2 = \frac{b-1}{bk-I}(V_b - V_r) \quad V_r = \sigma^2$$

$$V_r = QM_{RESÍDUO} = \sigma^2$$

$$V_b = QM_{BLOCOS(AJ.)} \text{ (quadrado médio para blocos ajustados)}$$

$$a = \frac{\sigma^2}{\sigma^2 + k\sigma_b^2} = \frac{V_r}{V_r + k \frac{b-1}{bk-I}(V_b - V_r)} = \frac{(bk-I)V_r}{k(b-1)V_b - (I-k)V_r}$$

Como $bk = JI$, então:

$$a = \frac{I(J-1)V_r}{k(b-1)V_b - (I-k)V_r}$$

b) Calcular, a partir do valor de $(\hat{a})$, as novas estimativas para os efeitos ajustados de tratamentos $(\hat{t}'_i)$ e para as suas médias ajustadas $(\hat{m}'_i)$, isto é, $\hat{m}'_i = \hat{m} + \hat{t}'_i$, tal que:

$$M_i = Q_i + \hat{a}A_i - \hat{a} \frac{k}{I} G$$

Em que G é o total geral, de modo que:

$$\hat{t}'_i = \frac{M_i}{\lambda I + (J - \lambda)\hat{a}}$$

Se $a = 0$, então $M_i = Q_i$. Logo, a análise corresponderá a intrablocos.

Dado que: $Q_i = kT_i - A_i$, então:

$$M_i = kT_i - A_i + \hat{a}A_i - \hat{a} \frac{k}{I} G$$

$$M_i = kT_i - (1 - \hat{a})A_i - \hat{a} \frac{k}{I} G$$

Se $\sigma_b^2 = 0$, então $a = 1$, logo:

$$M_i = kT_i - \frac{k}{I} G \quad \hat{t}'_i = \frac{M_i}{\lambda I + (J - \lambda)} = \frac{M_i}{\lambda(I - 1) + J}$$

Como nos blocos incompletos equilibrados, $\lambda(I - 1) = j(k - 1)$, então:

$$\hat{t}'_i = \frac{M_i}{\lambda(I - 1) + J} = \frac{M_i}{J(k - 1) + J} = \frac{M_i}{Jk} = \frac{1}{Jk} (kT_i - \frac{k}{I} G) = \frac{T_i}{J} - \frac{G}{IJ} = \frac{T_i}{J} - \hat{m}$$

Logo:

$$\hat{m}'_i = \hat{m} + \hat{t}'_i = \frac{T_i}{J}$$

Pimentel Gomes (2000) comenta que se $a = 1$, a análise procede como se não houvesse blocos incompletos. Em outras palavras, a análise recai na análise de um experimento casualizado em blocos, no caso de se tratar de um experimento do tipo I. Por outro lado, recai em um experimento inteiramente ao acaso se o delineamento for do tipo III.

A ocorrência de $\hat{V}_b < \hat{V}_r$ implica $\hat{\sigma}_b^2 < 0$, então se deve considerar $\hat{\sigma}_b^2 = 0$ e $\hat{a} = 1$. Por outro lado, se $\hat{V}_b$ apresentar um valor muito grande em relação à $\hat{V}_r = \hat{\sigma}^2$, o valor de $\hat{a}$ será praticamente nulo. Entretanto, se $\hat{a} = 0$, então $M_i = Q_i$, depreendendo que a análise com recuperação da informação interblocos equivale à análise intrablocos.

As estimativas de σ^2 e V_b são obtidas, respectivamente, pelo quadrado médio do resíduo ($\hat{\sigma}^2 = QM_{RESIDUO}$) e quadrado médio para blocos ajustados ($V_b = QM_{BLOCO(Aj.)}$).

c) Testar a significância para tratamentos ajustados.

Na análise com recuperação da informação interblocos, mediante o uso das estimativas de $\hat{\sigma}^2$ e $\hat{V}_b$, não existe um teste exato de significância para tratamentos. Entretanto, um teste aproximado de significância, para a hipótese de nulidade de efeitos de tratamentos, poderia ser realizado pelo teste de χ^2 , em que a expressão 9.3.2.1 é aproximadamente distribuída como um χ^2 com $I-1$ graus de liberdade.

$$\chi^2 = \frac{\sum_{i=1}^I M_i^2}{k\sigma^2[\lambda I + (J - \lambda)\hat{a}]} \quad 9.3.2.1$$

Por outro lado, Pimentel Gomes (2000) apresenta uma alternativa, que seria obter a soma de quadrados para tratamentos (ajustados), usando a expressão:

$$SQ_{TRAT(Aj.)}^* = \frac{\sum_{i=1}^I M_i^2}{k[\lambda I + (J - \lambda)\hat{a}]}$$

Resultando, pois, no cálculo do novo quadrado médio para tratamentos ajustados:

$$QM_{TRAT(Aj.)}^* = \frac{SQ_{TRAT(Aj.)}^*}{I-1} \rightarrow F = \frac{QM_{TRAT(Aj.)}^*}{\hat{\sigma}^2}$$

O uso do asterisco (*) objetiva diferenciar as expressões análogas às definidas na análise intrablocos, isto é, sem efetuar a recuperação da informação interblocos.

Pimentel Gomes (2000) adverte que os quadrados médios devem ter pelo menos vinte graus de liberdade para que a aproximação ao qui-quadrado χ^2 seja adequada.

d) Testar comparações múltiplas

Para os testes de comparações múltiplas de “*t*” de Student, de Tukey, de Bonferroni, de Duncan etc, tem-se:

$$\hat{V}(t'_i - t'_u) = V(\hat{m}'_i - \hat{m}'_u) = \frac{2k\hat{\sigma}^2}{\lambda I + (J - \lambda)\hat{a}} \quad (i \neq u)$$

No caso do contraste de tratamentos (Y):

$$Y = c_1 m_1 + c_2 m_2 + \dots + c_I m_I \text{ ou } Y = c_1 t_1 + c_2 t_2 + \dots + c_I t_I$$

$$\hat{V}(\hat{Y}) = \frac{2\hat{\sigma}^2}{\lambda I + (J - \lambda)\hat{a}} (c_1^2 + c_2^2 + \dots + c_I^2)$$

Exercício 9.3.2.1. Tipo I. Considerar um experimento, com uma espécie florestal (plano 9.1), que consiste em oito métodos de manejo florestal ($I = 8$) M1, M2, M3, M4, M5, M6, M7, M8. O delineamento utilizado foi de blocos incompletos equilibrados do tipo I, utilizando vinte e oito blocos ($b = 28$) — em que cada bloco apresenta duas parcelas ($k = 2$) de meio hectare —, distribuídos em sete repetições ($J = 7$). A variável resposta medida (dados simulados) foi o volume de madeira no final da rotação de dez anos. Observa-se, na configuração do ensaio, que qualquer tratamento aparece uma única vez ($\lambda = 1$) no mesmo bloco com cada um dos demais. Nesse contexto, efetuar a análise com a recuperação da informação interblocos.

Tabela 9.3.2.1. Dados do delineamento em bloco incompleto balanceado do tipo I.

Rep1		Total	Rep2		Total	Rep3		Total
M1=36,6	M2=32,9	69,5	M1=43,9	M3=32,9	76,8	M1=42,0	M4=31,1	73,1
M3=27,4	M4=29,2	56,6	M2=45,7	M8=29,7	75,4	M2=47,5	M7=32,9	80,4
M5=25,6	M6=27,4	53,0	M4=23,7	M5=29,2	52,9	M3=27,4	M6=31,1	58,5
M7=29,2	M8=27,9	57,1	M6=21,9	M7=29,2	51,1	M5=23,7	M8=24,2	47,9
		236,2			256,2			259,9
Rep4		Total	Rep5		Total	Rep6		Total
M1=38,4	M5=23,7	62,1	M1=51,2	M6=25,6	76,8	M1=40,2	M7=31,1	71,3
M2=42,0	M3=29,2	71,2	M2=49,4	M4=32,9	82,3	M2=43,9	M6=29,2	73,1
M4=18,3	M7=21,9	40,2	M3=32,9	M8=22,4	55,3	M3=32,9	M5=25,6	58,5
M6=23,7	M8=15,1	38,8	M5=29,2	M7=31,1	60,3	M4=27,4	M8=26,1	53,5
		212,3			274,7			256,4
Rep7		Total						
M1=42,0	M8=22,4	64,4						
M2=38,4	M5=23,7	62,1						
M3=27,4	M7=21,9	49,3						
M4=23,7	M6=29,2	52,9						
		228,7						

$$\lambda = \frac{J(k-1)}{I-1} = \frac{7(2-1)}{8-1} = 1$$

a) Somas de quadrados:

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 56.664,72 - \frac{1.724,4^2}{8 \times 7} = 3.565,5171$$

$$SQ_{REP} = \frac{1}{8}(236,2^2 + 256,2^2 + 259,9^2 + 212,3^2 + 274,7^2 + 256,4^2 + 228,7^2) - \frac{1.724,4^2}{56}$$

$$SQ_{REP} = \frac{1}{8}(427.552,92) - \frac{1.724,4^2}{56} = 344,9121$$

$$SQ_{TRAT(N\tilde{a}o Aj.)} = \frac{1}{J}(T_1^2 + T_2^2 + \dots + T_I^2) - C$$

$$SQ_{TRAT(N\tilde{a}o Aj.)} = \frac{1}{7}(390.460,5) - \frac{1.724,4^2}{56} = 2.680,8686$$

b) Soma de quadrados de blocos dentro das repetições (ajustadas) $SQ_{BLOCOS D. REP(Aj.)}$

$$SQ_{BLOCOS D. REP(Aj.)} = SQ_{BLOCOS(N\tilde{a}o Aj.)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(N\tilde{a}o Aj.)}$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{k}(B_1^2 + B_2^2 + \dots + B_b^2) - C$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{2}(109.978,18) - \frac{1.724,4^2}{56} = 1.889,8871$$

$$SQ_{BLOCOS D. REP(N\tilde{a}o Aj.)} = SQ_{BLOCOS(N\tilde{a}o Aj.)} - SQ_{REP} = 1.889,8871 - 344,9121 = 1.544,9750$$

$$SQ_{TRAT(Aj.)} = \frac{1}{\lambda k I} \sum_{i=1}^I Q_i^2 = \frac{1}{1 \times 2 \times 8} (23.317,1408) = 1.457,3213$$

$$SQ_{BLOCOS D. REP(Aj.)} = SQ_{BLOCOS D. REP(N\tilde{a}o Aj.)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(N\tilde{a}o Aj.)}$$

$$SQ_{BLOCOS D. REP(Aj.)} = 1.544,9750 + 1.457,3213 - 2.680,8686 = 321,4277$$

Tabela 9.3.2.2. ANOVA para os blocos ajustados dentro das repetições.

F.V.	G.L.	S.Q.	Q.M.	F
Repetições (Rep)	6	344,9121	57,4854	5,53**
Blocos d. Rep. (Aj.)	21	321,4277	$\hat{V}_b = 15,3061$	1,47 ns
Trat. (N\tilde{a}o Aj.)	7	2.680,8686	382,9812	36,84**
Resíduo	21	218,3087	$\hat{\sigma}^2 = 10,3957$	
Total	55	3.565,5171		

$$F_{6; 21; 0,01} = 3,81 \quad F_{7; 21; 0,01} = 3,64 \quad F_{21; 21; 0,05} = 2,09$$

c) Obtenção do valor de $(\hat{a})$, em que, nos delineamentos em blocos equilibrados do Tipo I:

$$\hat{a} = \frac{(J-1)\hat{\sigma}^2}{J\hat{V}_b - \hat{\sigma}^2} = \frac{6 \times 10,3957}{7 \times 15,3061 - 10,3957} = \frac{62,3742}{96,7470} = 0,6447$$

d) Cálculos dos valores de M_i , $\hat{t}'_i$ e $\hat{m}'_i = \hat{m} + \hat{t}'_i$:

$$M_i = Q_i + \hat{a}A_i - \hat{a}\frac{k}{I}G \quad \hat{t}'_i = \frac{M_i}{\lambda I + (J-\lambda)\hat{a}}$$

$$M_i = Q_i + 0,6467A_i - 0,6467 \times \frac{2}{8} \times 1.724,4$$

$$M_i = Q_i + 0,6467A_i - 278,7924$$

$$\hat{t}'_i = \frac{M_i}{\lambda I + (J-\lambda)\hat{a}} = \frac{M_i}{1 \times 8 + (7-1)0,6467} \rightarrow \hat{t}'_i = \frac{M_i}{11,8802}$$

$$\hat{m} = \frac{1.724,4}{56} = 30,7929 \quad \hat{m}'_i = 30,7929 + \hat{t}'_i$$

Tabela 9.3.2.3. Resultados preliminares para a análise de recuperação interblocos.

Trat.	T_i	A_i	$Q_i = kT_i - A_i$	M_i	$\hat{t}'_i$	$\hat{m}'_i = \hat{m} + \hat{t}'_i$
M1	294,3	494,0	94,6	135,277	11,3868	42,1797
M2	299,8	514,0	85,6	139,211	11,7179	42,5108
M3	210,1	426,2	-6,0	-9,169	-0,7718	30,0211
M4	186,3	411,5	-38,9	-51,575	-4,3413	26,4516
M5	180,7	396,8	-35,4	-57,582	-4,8469	25,9460
M6	188,1	404,2	-28,0	-45,396	-3,8212	26,9717
M7	197,3	409,7	-15,1	-28,939	-2,4359	28,3570
M8	167,8	392,4	-56,8	-81,827	-6,8877	23,9052
			0	0	0	

e) Teste aproximado para a hipótese de nulidade de efeitos de tratamentos.

Na análise interblocos com recuperação da informação não existe um teste exato de significância para os tratamentos. Dessa forma, deve-se adotar o seguinte procedimento aproximado para testar a hipótese de nulidade de efeito dos tratamentos, obtendo assim a soma de quadrados para tratamentos (ajustados) por meio da expressão:

$$SQ_{TRAT(Aj)}^* = \frac{\sum_{i=1}^I M_i^2}{k[\lambda I + (J-\lambda)\hat{a}]} = \frac{53.333,6}{2[1 \times 8 + (7-1) \times 0,6467]} = \frac{53.333,6}{23,7604} = 2.244,6423$$

$$QM_{TRAT(Aj)}^* = \frac{SQ_{TRAT(Aj)}^*}{I-1}$$

$$QM_{TRAT(Aj)}^* = \frac{2.244,6423}{7} = 320,6632$$

$$F_{TRAT(Aj)}^* = \frac{QM_{TRAT(Aj)}^*}{\hat{\sigma}^2} = \frac{320,6632}{10,3957} = 30,85^{**}$$

O asterisco é para diferenciar das expressões análogas as definidas na análise intrablocos.

f) Teste de comparações múltiplas de Tukey

$$\hat{V}(\hat{t}'_i - \hat{t}'_u) = \hat{V}(\hat{m}'_i - \hat{m}'_u) = \frac{2k\hat{\sigma}^2}{\lambda I + (J - \lambda)\hat{a}}$$

$$\hat{V}(\hat{t}'_i - \hat{t}'_u) = \hat{V}(\hat{m}'_i - \hat{m}'_u) = \frac{2 \times 2 \times 10,3957}{1 \times 8 + (7 - 1) \times 0,6467} = \frac{41,5828}{11,8802} = 3,5002$$

$$\Delta = q \sqrt{\frac{1}{2} \hat{V}(\hat{t}'_i - \hat{t}'_u)} = 4,74 \sqrt{\frac{3,5002}{2}} = 6,27$$

$$q_{8; 21; 0,05} \cong 4,74$$

Tratamentos	Médias Ajustadas
M2	42,51 a
M1	42,18 a
M3	30,02 b
M7	28,36 b
M6	26,97 b
M5	25,94 b
M4	26,45 b
M8	23,90 b

Médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem entre si pelo teste de Tukey. As médias dos manejos M2 e M1 apresentaram os melhores desempenhos e não diferem significativamente entre si.

Exercício 9.3.2.2. Tipo II. Considerar um experimento (plano 9.3) com sete ($I = 7$) métodos de manejo de uma espécie florestal ($M1, M2, \dots, M7$). O delineamento foi de blocos incompletos equilibrados do tipo II com vinte e um blocos ($b = 21$), apresentando duas parcelas ($k = 2$) de meio hectare cada. A variável resposta (dados simulados) foi o volume de madeira no final da rotação. No ensaio, observa-se que qualquer tratamento aparece uma única vez ($\lambda = 1$), no mesmo bloco, com cada um dos demais. O número de repetições é igual a seis ($J = 6$). Efetuar a análise com recuperação da informação interblocos.

Tabela 9.3.2.4. Dados do delineamento em bloco incompleto equilibrados do tipo II.

GRUPO1		Totais	GRUPO2		Totais	GRUPO3		Totais
M1=81,5	M2=65,2	146,7	M1=88,5	M3=93,2	181,7	M1=69,9	M4=51,2	121,1
M2=74,5	M3=86,2	160,7	M3=83,8	M5=55,9	139,7	M4=62,9	M7=79,2	142,1
M3=81,5	M4=58,2	139,7	M5=46,6	M7=69,9	116,5	M7=74,5	M3=90,8	165,3
M4=65,2	M5=55,9	121,1	M7=65,2	M2=58,2	123,4	M3=76,8	M6=55,9	132,7
M5=62,9	M6=74,5	137,4	M2=60,5	M4=65,2	125,7	M6=65,2	M2=79,2	144,4
M6=55,9	M7=60,5	116,4	M4=53,5	M6=55,9	109,4	M2=67,5	M5=53,5	121,0
M7=72,2	M1=62,9	135,1	M6=65,2	M1=76,8	142,0	M5=46,6	M1=76,8	123,4
		957,1			938,4			950,0

a) Eficiência (E) e λ :

$$\lambda = \frac{J(k-1)}{I-1} = \frac{6(2-1)}{7-1} = 1$$

$$E = \frac{\lambda I}{kJ} = \frac{1 \times 7}{2 \times 6} = 0,5833 \rightarrow 58,33\%$$

b) Somas de quadrados:

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 198.715,93 - \frac{2.845,5^2}{7 \times 6} = 5.933,3050$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{k} (B_1^2 + B_2^2 + \dots + B_b^2) - C$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{2} (146,7^2 + 160,7^2 + \dots + 121,0^2 + 123,4^2) - \frac{2.845,5^2}{42}$$

$$SQ_{BLOCOS(N\tilde{a}o Aj.)} = \frac{1}{2} (391.980,97) - \frac{2.845,5^2}{42} = 3.207,86$$

$$SQ_{GRUPOS} = \frac{1}{14} (957,1^2 + 938,4^2 + 950,0^2) - \frac{2.845,5^2}{42}$$

$$SQ_{GRUPOS} = \frac{1}{14} (2.699.134,97) - \frac{2.845,5^2}{42} = 12,73$$

$$SQ_{BLOCOS D. GRUPOS(N\tilde{a}o Aj.)} = SQ_{BLOCOS(N\tilde{a}o Aj.)} - SQ_{GRUPOS} = 3.207,86 - 12,73 = 3.195,13$$

$$SQ_{TRAT(N\tilde{a}o Aj.)} = \frac{1}{6} (456,4^2 + 405,1^2 + 512,3^2 + 356,2^2 + 321,4^2 + 372,6^2 + 421,5^2) - \frac{2.845,5^2}{42}$$

$$SQ_{TRAT(Não Aj.)} = \frac{1}{6}(1.181.527,7) - \frac{2.845,5^2}{42} = 4.138,6583$$

c) Para obter o valor da soma de quadrados para tratamentos ajustados para os efeitos de blocos $SQ_{TRAT(Aj.)}$, é necessário calcular:

c1) os totais por tratamentos T_i ($i=1, \dots, I$)

T1	456,4	T3	512,3	T5	321,4	T7	421,5
T2	405,1	T4	356,2	T6	372,6		

c2) a soma total dos blocos em que aparece o tratamento i correspondente (A_i):

A1	146,7+181,7+121,1+135,1+142,0+123,4	850,0
A2	146,7+160,7+123,4+125,7+144,4+121,0	821,9
A3	181,7+160,7+139,7+139,7+165,3+132,7	919,8
A4	121,1+142,1+139,7+121,1+125,7+109,4	759,1
A5	139,7+116,5+121,1+137,4+121,0+123,4	759,1
A6	132,7+137,4+144,4+116,4+109,4+142,0	782,3
A7	142,1+116,5+165,3+123,4+116,4+135,1	798,8

c3) calcular $Q_i = kT_i - A_i$

A soma de quadrados de tratamentos ajustados é obtida a partir dos dados da tabela 9.3.2.6:

$$SQ_{TRAT(Aj.)} = \frac{\sum_{i=1}^I Q_i^2}{\lambda k I} = \frac{34.100,4}{1 \times 2 \times 7} = 2.435,7429$$

$$SQ_{BLOCOS D. GRUPO(Aj)} = SQ_{BLOCOS D. GRUPO(Não Aj)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(Não Aj)}$$

$$SQ_{BLOCOS D. GRUPO(Aj)} = 3.195,13 + 2.435,7429 - 4.138,6583 = 1.492,2146$$

Tabela 9.3.2.5. ANOVA, considerando blocos dentro dos grupos ajustados.

F.V.	G.L.	S.Q.	Q.M.	F
Grupos	2	12,7300	6,3650	0,33 <i>ns</i>
Blocos d. Grupo (Aj.)	18	1.492,2146	$\hat{\nu}_b = 82,9008$	4,29**
Trat. (Não Aj)	6	4.138,6583	689,7764	
Resíduo	15	289,7021	$\hat{\sigma}^2 = 19,3135$	
Total	41	5.933,3050		

$$F_{2;15;0,05} = 3,68 \quad F_{18;15;0,01} = 2,36 \quad F_{6;15;0,01} = 4,32$$

d) Obter o valor de ($\hat{a}$)

Para os delineamentos em blocos equilibrados do Tipo II, tem-se:

$$\hat{a} = \frac{[I(J-1) - k(g-1)]\hat{\sigma}^2}{k(b-g)\hat{V}_b - (I-k)\hat{\sigma}^2} = \frac{[7(6-1) - 2(3-1)] \times 19,3135}{2(21-3) \times 82,9008 - (7-2) \times 19,3135} = \frac{598,7185}{2.887,8613} = 0,2073$$

e) Calcular os valores de M_i , $\hat{t}'_i$ e $\hat{m}'_i = \hat{m} + \hat{t}'_i$:

$$M_i = Q_i + \hat{a}A_i - \hat{a} \frac{k}{I} G = Q_i + 0,2073 A_i - 0,2073 \times \frac{2}{7} \times 2.845,5$$

$$M_i = Q_i + 0,2073 A_i - 168,5349$$

$$\hat{t}'_i = \frac{M_i}{\lambda I + (J - \lambda)\hat{a}} \quad \hat{t}'_i = \frac{M_i}{1 \times 7 + (6 - 1) \times 0,2073} \quad \hat{t}'_i = \frac{M_i}{8,0365}$$

$$\hat{m} = \frac{\sum_{i=1}^I \sum_{j=1}^J x_{ij}}{IJ} = \frac{2.845,5}{42} = 59,2813 \quad \hat{m}'_i = 59,2813 + \hat{t}'_i$$

Os valores de $\hat{t}'_i$ e $\hat{m}'_i$ estão apresentados na tabela 9.3.2.6.

Tabela 9.3.2.6. Resultados dos cálculos para a análise de recuperação interblocos.

Trat.	T_i	A_i	$Q_i = kT_i - A_i$	M_i	$\hat{t}'_i = \frac{M_i}{\lambda I + (J - \lambda)\hat{a}}$	$\hat{m}'_i = \hat{m} + \hat{t}'_i$
M1	456,4	850,0	62,8	70,470	8,7688	68,0501
M2	405,1	821,9	-11,7	-9,855	-1,2263	58,0550
M3	512,3	919,8	104,8	126,940	15,7954	75,0767
M4	356,2	759,1	-46,7	-57,873	-7,2013	52,0800
M5	321,4	759,1	-116,3	-127,473	-15,8618	43,4195
M6	372,6	782,3	-37,1	-43,464	-5,4083	53,8730
M7	421,5	798,8	44,2	41,256	5,1336	64,4149
			0	0	0	

f) Teste aproximado para a hipótese de nulidade de efeitos de tratamentos.

Na análise com recuperação da informação interblocos não existe um teste exato de significância para os tratamentos. Considerar o seguinte procedimento aproximado para testar a hipótese de nulidade de efeito dos tratamentos e assim obter a soma de quadrados para tratamentos (ajustados) por meio da expressão:

$$SQ_{TRAT(Aj)}^* = \frac{\sum_{i=1}^I M_i^2}{k[\lambda I + (J - \lambda)]\hat{a}} = \frac{44.366,9}{2[1 \times 7 + (6 - 1) \times 0,2073]}$$

$$SQ_{TRAT(Aj)}^* = \frac{\sum_{i=1}^I M_i^2}{k[\lambda I + (J - \lambda)]\hat{a}} = \frac{44.366,9}{16,073} = 2.760,3372$$

$$QM_{TRAT(Aj)}^* = \frac{QM_{TRAT(Aj)}^*}{(I - 1)} = \frac{2.760,3372}{6} = 460,0562$$

$$F = \frac{QM_{TRAT(Aj)}^*}{\hat{\sigma}^2} = \frac{460,0562}{19,3135} = 23,82^{**}$$

O uso do asterisco objetiva apenas diferenciar as expressões análogas definidas na análise intra-blocos.

g) Teste de comparação múltipla de Tukey

$$\hat{V}(\hat{m}'_i - \hat{m}'_u) = \frac{2k\hat{\sigma}^2}{\lambda I + (J - \lambda)\hat{a}} = \frac{2 \times 2 \times 19,3135}{1 \times 7 + (6 - 1) \times 0,2073} = \frac{77,2540}{8,0365} = 9,6129$$

$$\Delta = q_{(7; 15; 0,05)} \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}'_i - \hat{m}'_u)} \quad \Delta = 4,78 \sqrt{\left(\frac{1}{2}\right) 9,6129} = 10,48$$

É importante observar que, conforme o exercício 9.3.1.2, o qual foi analisado pelo método intrablocos, o valor da variância foi $\hat{V}(\hat{m}'_i - \hat{m}'_u) = 11,0363$, enquanto a análise com recuperação da informação interblocos apresentou o valor de $\hat{V}(\hat{m}'_i - \hat{m}'_u) = 9,6129$, mostrando que as médias estimadas pelo método com recuperação da informação interblocos são comparadas com maior precisão.

Tratamentos	Médias Ajustadas
M3	75,08 a
M1	68,05 ab
M7	64,41 b
M2	58,06 bc
M6	53,87 cd
M4	52,08 cd
M5	43,42 d

Médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem entre si pelo teste de Tukey. As médias dos manejos M3 e M1 apresentaram os melhores desempenhos e não diferem significativamente.

Exercício 9.3.2.3. Tipo III. De acordo com o plano 9.6, considerar um experimento de competição de treze ($I = 13$) procedências ($P1, P2, \dots, P13$) de uma espécie florestal. O delineamento utilizado foi de blocos incompletos equilibrados do tipo III, considerando treze blocos ($b = 13$) com quatro parcelas ($k = 4$) de meio hectare cada e quatro repetições ($J = 4$). A variável resposta medida foi o volume de madeira no final da rotação de dez anos. Efetuar a análise pelo método de recuperação da informação interblocos. Na tabela 9.3.2.7, estão os dados (simulados) para análise de recuperação interblocos do experimento tipo III.

$$\lambda = \frac{J(k-1)}{I-1} = \frac{4(4-1)}{13-1} = 1$$

Tabela 9.3.2.7. Dados para análise de recuperação interblocos experimento tipo III.

Blocos	Volume de madeira por procedência (m³ por 0,5 ha)								Total
1	P3	54,0	P6	42,4	P9	61,9	P11	42,5	200,8
2	P3	49,1	P4	42,2	P8	76,1	P12	48,4	215,8
3	P10	34,0	P11	31,2	P12	67,6	P13	66,8	199,6
4	P2	58,3	P5	57,6	P8	81,0	P11	27,1	224,0
5	P7	49,9	P8	70,1	P9	65,7	P10	69,1	254,8
6	P4	65,3	P5	69,1	P6	58,0	P10	70,0	262,4
7	P1	74,1	P5	66,4	P9	54,8	P12	65,1	260,4
8	P3	73,4	P5	69,1	P7	71,1	P13	88,8	302,4
9	P1	81,5	P2	70,2	P3	79,6	P10	66,8	298,1
10	P2	61,2	P4	65,5	P9	57,4	P13	85,3	269,4
11	P1	78,1	P4	66,4	P7	66,4	P11	50,6	261,5
12	P1	67,9	P6	71,9	P8	64,3	P13	97,7	301,8
13	P2	64,7	P6	67,2	P7	83,9	P12	57,0	272,8

a) Cálculo das somas de quadrados

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 222.795,66 - \frac{3.323,8^2}{13 \times 4} = 10.340,9208$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{4} (200,8^2 + 215,8^2 + \dots + 301,8^2 + 272,8^2) - \frac{3.323,8^2}{52} = 3.860,8758$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{4} (865.262,46) - \frac{3.323,8^2}{52} = 3.860,8758$$

$$SQ_{TRAT (N\tilde{a}o Aj.)} = \frac{1}{J} (T_1^2 + T_2^2 + \dots + T_I^2) - C$$

$$SQ_{TRAT (N\tilde{a}o Aj.)} = \frac{1}{4} (301,6^2 + 254,4^2 + \dots + 238,1^2 + 338,6^2) - C$$

$$SQ_{TRAT(N\tilde{a}o Aj.)} = \frac{1}{4}(872.586,1) - \frac{3.323,8^2}{52} = 5.691,7858$$

b) Soma de quadrados para tratamentos ajustados $SQ_{TRAT(Aj.)}$ para os efeitos de blocos:

b1) obter os totais por tratamentos $[T_i \ (i=1,\dots,I)]$ que estão mostrados na tabela 9.3.2.10.

b2) calcular a soma total dos blocos em que aparece o tratamento i correspondente (A_i):

		A_i
P1	260,4+298,1+261,5+301,8	1.121,8
P2	224,0+269,4+272,8+298,1	1.064,3
P3	200,8+215,8+302,4+298,1	1.017,1
P4	262,4+215,8+269,4+261,5	1.009,1
P5	224,0+262,4+260,4+302,4	1.049,2
P6	200,8+301,8+272,8+262,4	1.037,8
P7	254,8+302,4+261,5+272,8	1.091,5
P8	254,8+215,8+224,0+301,8	996,4
P9	200,8+254,8+260,4+269,4	985,4
P10	199,6+254,8+262,4+298,1	1.014,9
P11	199,6+200,8+224,0+261,5	885,9
P12	199,6+215,8+260,4+272,8	948,6
P13	199,6+302,4+269,4+301,8	1.073,2

b3) obter $Q_i = kT_i - A_i$

A soma de quadrados dos tratamentos ajustados (valores de Q_i na tabela 9.3.2.10).

$$SQ_{TRAT(Aj.)} = \frac{1}{\lambda KI} \sum_{i=1}^I Q_i^2 = \frac{1}{1 \times 4 \times 13} 208.616,14 = 4.011,8488$$

Então, pode-se calcular a soma de quadrados dos blocos ajustados:

$$SQ_{BLOCOS(Aj.)} = SQ_{BLOCOS(N\tilde{a}o Aj.)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(N\tilde{a}o Aj.)}$$

$$SQ_{BLOCOS(Aj.)} = 3.860,8758 + 4.011,8488 - 5.691,7858 = 2.180,9388$$

Tabela 9.3.2.8. ANOVA com blocos não ajustados (BIB tipo III).

F.V.	G.L.	S.Q.	Q.M.	F
Blocos (Não Aj.)	12	3.860,8758	321,7396	3,52**
Trat. (Aj)	12	4.011,8488	334,3207	3,66**
Resíduo	27	2.468,1962	91,4147	
Total	51	10.340,9208		

Tabela 9.3.2.9. ANOVA com blocos ajustados (BIB tipo III).

F.V.	G.L.	S.Q.	Q.M.	F
Blocos (Aj.)	12	2.180,9388	$\hat{V}_b = 181,7449$	1,99 <i>ns</i>
Trat. (Não Aj)	12	5.691,7858	474,3154	5,1886
Resíduo	27	2.468,1962	$\hat{\sigma}^2 = 91,4147$	
Total	51	10.340,9208		

$$F_{12; 27; 0,01} = 2,93$$

c) Estimar o valor de ($\hat{a}$)

Nos delineamentos em blocos equilibrados do Tipo III, tem-se:

$$\hat{a} = \frac{I(J-1)\hat{\sigma}^2}{k(b-1)\hat{V}_b - (I-k)\hat{\sigma}^2} = \frac{13(4-1) \times 91,4147}{4(13-1) \times 181,7449 - (13-4) \times 91,4147}$$

$$\hat{a} = \frac{I(J-1)\hat{\sigma}^2}{k(b-1)\hat{V}_b - (I-k)\hat{\sigma}^2} = \frac{3.565,1733}{7.901,0229} = 0,451230$$

d) Obter os valores de M_i , $\hat{t}'_i$ e $\hat{m}'_i$

Tabela 9.3.2.10. Valores de M_i , $\hat{t}'_i$ e $\hat{m}'_i$ para a análise de recuperação interblocos.

	T_i	A_i	$Q_i = kT_i - A_i$	M_i	$\hat{t}'_i$	$\hat{m}'_i = \hat{m} + \hat{t}'_i$
P1	301,6	1121,8	84,6	129,313	9,0091	72,9283
P2	254,4	1064,3	-46,7	-27,932	-1,9460	61,9732
P3	256,1	1017,1	7,3	4,770	0,3323	64,2515
P4	239,4	1009,1	-51,5	-57,640	-4,0157	59,9035
P5	262,2	1049,2	-0,4	11,554	0,8050	64,7242
P6	239,5	1037,8	-79,8	-72,990	-5,0851	58,8341
P7	271,3	1091,5	-6,3	24,741	1,7237	65,6429
P8	291,5	996,4	169,6	157,729	10,9888	74,9080
P9	239,8	985,4	-26,2	-43,034	-2,9981	60,9211
P10	239,9	1014,9	-55,3	-58,823	-4,0981	59,8211
P11	151,4	885,9	-280,3	-342,032	-23,8288	40,0904
P12	238,1	948,6	3,8	-29,640	-2,0649	61,8543
P13	338,6	1073,2	281,2	303,984	21,1781	85,0973
			0	0	0	

Tal que:

$$M_i = Q_i + \hat{a}A_i - \hat{a} \frac{k}{I} G \quad \hat{t}'_i = \frac{M_i}{\lambda I + (J - \lambda)\hat{a}}$$

$$M_i = Q_i + 0,451230A_i - 0,451230 \times \frac{4}{13} \times 3.323,8 \quad M_i = Q_i + 0,451230A_i - 461,476392$$

$$\hat{t}'_i = \frac{M_i}{1 \times 13 + (4-1)0,451230} \quad \hat{t}'_i = \frac{M_i}{14,353690}$$

$$\hat{m} = \frac{3.323,8}{52} = 63,9192 \quad \hat{m}'_i = 63,9192 + \hat{t}'_i$$

e) Teste aproximado para a hipótese de nulidade de efeitos de tratamentos.

Na análise interblocos com recuperação da informação não existe um teste exato de significância para os tratamentos, então se adota o seguinte procedimento aproximado para testar a hipótese de nulidade de efeito dos tratamentos, obtendo-se a soma de quadrados para tratamentos (ajustados) por meio da expressão:

$$SQ_{TRAT(Aj)}^* = \frac{\sum_{i=1}^I M_i^2}{k[\lambda I + (J - \lambda)\hat{a}]} = \frac{267.381,3}{4[(1 \times 13) + (4 - 1) \times 0,451230]} = \frac{267.381,3}{57,414760} = 4.657,0133$$

$$QM_{TRAT(Aj)}^* = \frac{4.657,0133}{12} = 388,0844$$

$$F_{TRAT(Aj)} = \frac{QM_{TRAT(Aj)}^*}{\hat{\sigma}^2} = \frac{388,0844}{91,4147} = 4,25^{**}$$

$$F_{12; 27; 0,01} = 2,88$$

O uso do asterisco objetiva apenas diferenciar das expressões análogas definidas na análise intra-blocos. Alternativamente, tem-se:

$$\chi_{I-1}^2 = \chi_{12}^2 = \frac{\sum_{i=1}^I M_i^2}{k\sigma^2[\lambda I + (J - \lambda)a]} = \frac{267.381,3}{4 \times 91,4147 \times [1 \times 13 + (4 - 1)0,451230]} = \frac{267.381,3}{5.248,5531} = 50,94^{**}$$

$$\chi_{12; 27; 0,01}^2 = 43,2$$

f) Teste de comparação múltipla de Tukey

$$\hat{V}(\hat{t}'_i - \hat{t}'_u) = \hat{V}(\hat{m}'_i - \hat{m}'_u) = \frac{2k\hat{\sigma}^2}{\lambda I + (J - \lambda)\hat{a}} = \frac{2 \times 4 \times 91,4147}{1 \times 13 + (4 - 1) \times 0,451230} = \frac{731,3176}{14,353690} = 50,9498$$

$$\Delta = q \sqrt{\frac{1}{2} \hat{V}(\hat{m}'_i - \hat{m}'_u)} = 5,12 \sqrt{\frac{50,9498}{2}} = 25,84$$

$$q_{13; 27; 0,05} \cong 5,12$$

É importante observar que, na análise intrablocos do exercício 9.3.1.3, o valor obtido da variância foi $\hat{V}(\hat{m}'_i - \hat{m}'_u) = 56,2552$, enquanto — considerando os mesmos dados, mas efetuando a análise com recuperação da informação interblocos — o valor calculado da variância foi de $\hat{V}(\hat{m}'_i - \hat{m}'_u) = 50,9498$, o que demonstra que, na análise com recuperação da informação interblocos, as médias estimadas são comparadas com maior precisão.

Tratamentos	Médias ajustadas
P13	85,10 a
P8	74,91 ab
P1	72,93 ab
P7	65,64 abc
P5	64,72 abc
P3	64,25 abc
P2	61,97 abc
P12	61,85 abc
P9	60,92 abc
P4	59,90 abc
P10	59,82 abc
P6	58,83 bc
P11	40,09 c

Médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem entre si pelo teste de Tukey. A procedência P13 apresentou o melhor desempenho.

9.4. BIBLIOGRAFIA

CAMPOS, H.; PIEDADE, S. M. S. **Planejamento de experimentos**. 1ª ed. Piracicaba: FEALQ, 1992. 185 p.

COCHRAN, W. G.; COX, G. M. **Experimental designs**. 2ª ed. New York: John Wiley, 1957. 611 p.

PIMENTEL GOMES, F.; GARCIA, C. H. **Estatística aplicada a experimentos agrônômicos e florestais**. Piracicaba: FEALQ. (Biblioteca de Ciências Agrárias Luiz de Queiroz - Volume 11), 2002. 309 p.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 14^a ed. São Paulo: Nobel, 2000. 475 p.

YATES, F. A new method of arranging variety trial involving a large number of varieties. **Jour. Agr. Sci.** 26,424-455, 1936.

YATES, F. The recovery of inter-block information in balanced incomplete block designs . *Annals of Eugenics*, Vol. 10, n.1, p. 317-325, 1940.

TREVISOL, A. M.; COSME, K. F. S. G. **Análise de experimentos em blocos incompletos**. Relatório apresentado à disciplina de estágio supervisionado II do curso de estatística da UNB, 2013. 69 p.

CAPÍTULO 10

EXPERIMENTOS RETICULADOS OU LÁTICES QUADRADOS

Este capítulo apresenta definições, conceitos e aplicações sobre experimentos reticulados ou em látices. Define os experimentos: látices quadrados (*square lattices*), látices retangulares (*rectangular lattices*), látices cúbicos (*cubic lattices*) e quadrados látices (*lattices square*). Enfatiza, com mais detalhes, os delineamentos látices quadrados que são os mais usados. Apresenta ainda exemplos de aplicação, utilizando os métodos de análise estatística intrablocos e o da análise com recuperação da informação interblocos. Utiliza a aplicação do teste de comparação múltipla de Tukey e mostra como calcular a eficiência, considerando as análises intrablocos e a recuperação da informação interblocos.

10.1. INTRODUÇÃO

Os delineamentos em látices são indicados quando o experimento apresenta um número muito grande de tratamentos, assim como são importantes para eliminar a heterogeneidade dentro dos blocos. Os delineamentos em látices são denominados incompletos, pois os tratamentos são arranjados, possibilitando que cada bloco contenha apenas parte dos tratamentos, assim como conjuntos de blocos podem originar grupos ou repetições com todos os tratamentos. De acordo com Pimentel Gomes (2000), os delineamentos em látices não são recomendados para experimentos de adubação e nem para ensaios fatoriais.

Os delineamentos em látices são classificados, quanto às comparações entre tratamentos em balanceados e parcialmente balanceados. São balanceados quando cada tratamento ocorre uma e só uma vez no mesmo bloco com todos os outros. Nesse caso, as comparações entre os tratamentos apresentam a mesma precisão, pois ocorrem simultaneamente em todos os blocos. Por outro lado, nos delineamentos parcialmente balanceados, isto é, quando apenas alguns tratamentos ocorrem juntos no mesmo bloco, haverá a necessidade de se calcular as diferenças entre as médias dos tratamentos que estão no mesmo bloco, assim como daquelas que não aparecem no mesmo bloco. Vale destacar que as comparações entre as médias dos tratamentos que ocorrem juntos, no mesmo bloco, tendem a apresentar maior precisão.

Os delineamentos reticulados quadrados (*Square Lattice*) foram introduzidos por Yates (1936). De acordo com Pimentel Gomes e Garcia (2002), os delineamentos reticulados (ou látices) apresentam os seguintes tipos básicos:

- Látices quadrados (*Square Lattices*);
- Látices retangulares (*Rectangular Lattices*);
- Látices cúbicos (*Cubic Lattices*);
- Quadrados látices (*Lattice Square*).

a) Látices quadrados são delineamentos que apresentam o número de tratamentos (I) igual a um quadrado perfeito ($I = k^2$). Por exemplo, $I = k^2 = 5^2 = 25$, em que os tratamentos são distribuídos em repetições e apresentando cada repetição k blocos com k parcelas. Esta obra abordará apenas os látices quadrados (ou reticulados quadrados) que são os mais utilizados. Para ilustrar, considerar o plano 10.1, apresentado em Cochran e Cox (1957), para o caso de um látice quadrado balanceado 3×3 estabelecido em doze blocos.

Plano 10.1. Látice quadrado balanceado com 3×3 .

Blocos	Rep1			Blocos	Rep2			Blocos	Rep3			Blocos	Rep4		
B1	A	B	C	B4	A	D	G	B7	A	E	I	B10	A	F	H
B2	D	E	F	B5	B	E	H	B8	B	F	G	B11	B	D	I
B3	G	H	I	B6	C	F	I	B9	C	D	H	B12	C	E	G

b) Látices retangulares são delineamentos que apresentam o número de tratamentos igual a $I = k \times (k + 1)$, em que cada bloco contém k parcelas. Exemplos: 3×4 , 7×8 etc.

Considerar o plano 10.2, apresentado em Cochran e Cox (1957), do látice 3×4 , com $J = 3$ repetições:

Plano 10.2. Delineamento látice retangular 3×4 .

Blocos	Rep1			Blocos	Rep2			Blocos	Rep3		
B1	A	B	C	B5	D	G	J	B9	F	H	L
B2	D	E	F	B6	A	H	K	B10	B	I	J
B3	G	H	I	B7	B	E	L	B11	C	D	K
B4	J	K	L	B8	C	F	I	B12	A	E	G

Tabela 10.1.1. ANOVA do látice retangular 3×4 .

F.V.	G.L.
Repetições	$J - 1 = 2$
Tratamentos	$k^2 + k - 1 = 11$
Blocos dentro das repetições	$Jk = 9$
Resíduo	$(J - 1)(k^2 - 1) - k = 13$
Total	$Jk^2 + Jk - 1 = 35$

c) Látices cúbicos são delineamentos em que o número de tratamentos é um cubo perfeito ($I = k^3$), em que o tamanho de cada bloco é a raiz cúbica do número de tratamentos, por exemplo: $I = 3^3$, $I = 4^3$ etc.

Cochran e Cox (1957) apresentam um procedimento para obter o planejamento para o delineamento $I = k^3 = 3^3$ (plano 10.3), no qual os tratamentos são numerados por meio de um código de três dígitos, sendo que cada dígito é constituído dos valores de 1 a k . Esses autores mostram um delineamento com 27 tratamentos, cada um com três repetições. Observa-se que cada repetição é composta por três subgrupos de $k^2 = 9$ tratamentos.

Plano 10.3. Delineamento látice cúbico ($I = 3^3 = 27$) tratamentos.

Repetição1																	
Bloco1		Bloco2		Bloco3		Bloco4		Bloco5		Bloco6		Bloco7		Bloco8		Bloco9	
T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod
1	111	4	121	7	131	10	112	13	122	16	132	19	113	22	123	25	133
2	211	5	221	8	231	11	212	14	222	17	232	20	213	23	223	26	233
3	311	6	321	9	331	12	312	15	322	18	332	21	313	24	323	27	333
Repetição2																	
Bloco1		Bloco2		Bloco3		Bloco4		Bloco5		Bloco6		Bloco7		Bloco8		Bloco9	
T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod
1	111	2	211	3	311	10	112	11	212	12	312	19	113	20	213	21	313
4	121	5	221	6	321	13	122	14	222	15	322	22	123	23	223	24	323
7	131	8	231	9	331	16	132	17	232	18	332	25	133	26	233	27	333
Repetição3																	
Bloco1		Bloco2		Bloco3		Bloco4		Bloco5		Bloco6		Bloco7		Bloco8		Bloco9	
T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod	T	Cod
1	111	2	211	3	311	4	121	5	221	6	321	7	131	8	231	9	331
10	112	11	212	12	312	13	122	14	222	15	322	16	132	17	232	18	332
19	113	20	213	21	313	22	123	23	223	24	323	25	133	26	233	27	333

Na primeira repetição, o dígito para os três primeiros tratamentos (1, 2 e 3) apresenta a seguinte construção: 11 para os dois últimos números, enquanto o primeiro representa os valores de $k = 1, 2, 3$, resultando na seguinte composição para o bloco 1 (111, 211 e 311). Para os próximos três tratamentos (4, 5 e 6), os dois últimos dígitos são representados por 21, gerando o bloco 2 (121, 221 e 321), assim sucessivamente, até o valor 33, resultando no bloco 9 (133, 233 e 333). No caso de $I = 3^3 = 27$, o tratamento 27 é representado por 333.

Na segunda repetição, o primeiro e o terceiro dígitos são fixos dentro dos blocos, enquanto o segundo dígito varia de 1 a k . No caso de 27 tratamentos, o primeiro bloco seria (111, 121 e 131), o segundo bloco (211, 221 e 231) e o nono bloco (313, 323 e 333).

Na terceira repetição, o primeiro e o segundo dígitos são constantes dentro de cada bloco. No delineamento látice cúbico com ($I = 3^3$) de 27 tratamentos (plano 10.3):

$$\lambda = \frac{J(k-1)}{I-1} = \frac{3(3-1)}{27-1} = \frac{6}{26}$$

No caso de um experimento balanceado $I = 3^3$ ($k=3$), em que $\lambda = 1$, seriam necessárias treze repetições ($J = 13$). Isto é:

$$\lambda = \frac{J(k-1)}{I-1} = \frac{13(3-1)}{27-1} = 1$$

d) Quadrados látices (*Lattice Square*) ou quadrados reticulados são delineamentos estruturados com o número de tratamentos (I) igual a um quadrado perfeito ($I = k^2$), mas dispostos em linhas e colunas semelhantes aos quadrados latinos com k parcelas cada. Considerar o plano 10.4 apresentado em Cochran e Cox (1957).

Plano 10.4. Quadrado látice balanceado 4×4 .

Quadrado I				Quadrado II				Quadrado III				Quadrado IV				Quadrado V			
A	E	I	M	A	B	C	D	A	K	P	F	A	G	L	N	A	J	O	H
B	F	J	N	F	E	H	G	L	B	E	O	H	B	M	K	I	B	G	P
C	G	K	O	K	L	I	J	N	H	C	I	J	P	C	E	M	F	C	L
D	H	L	P	P	O	N	M	G	M	J	D	O	I	F	D	E	N	K	D

Tabela 10.1.2. ANOVA do quadrado látice balanceado 4×4 .

F.V.	G.L.
Repetições	$k = 4$
Tratamentos	$k^2 - 1 = 15$
Linhas (ajustadas para tratamentos)	$k^2 - 1 = 15$
Linhas (ajustadas para tratamentos e colunas)	$k^2 - 1 = 15$
Colunas (ajustadas para tratamentos)	$k^2 - 1 = 15$
Colunas (ajustadas para tratamentos e linhas)	$k^2 - 1 = 15$
Resíduo	$(k^2 - 1)(k - 2) = 30$
Total	$k^3 + k^2 - 1 = 79$

$$\lambda = \frac{J(k-1)}{I-1} = \frac{5(4-1)}{16-1} = 1$$

De acordo com Pimentel Gomes (2000), os delineamentos látices retangulares, látices cúbicos e quadrados látices são pouco usados, enquanto os látices quadrados são os mais utilizados.

Moraes *et al.* (1988) comentam que um dos delineamentos em látices mais utilizados na pesquisa florestal é o quadrado parcialmente balanceado e triplo, sem repetição dos planos básicos. Os autores apresentam uma aplicação de um látice triplo 8×8 com três repetições. Para a análise de variância do experimento, foram utilizadas as médias de altura de um teste de progênie de *Eucalyptus grandis*.

Existem três procedimentos diferentes de análise estatística para os delineamentos látices:

- a) A análise mais simples consiste em considerar o ensaio como um experimento casualizado em blocos. Entretanto, isso é recomendável apenas quando os efeitos de blocos são pequenos ou nulos;
- b) A análise intrablocos é exata e leva em conta as diferenças entre blocos. A análise intrablocos dos látices se baseia no método dos quadrados mínimos, tal como a que é feita para os delineamentos inteiramente casualizados ou em blocos completos;
- c) A análise com recuperação da informação interblocos é a mais eficiente, entretanto, produz resultados apenas aproximados.

10.2. LÁTICES QUADRADOS (*SQUARE LATTICES*)

Látices quadrados são delineamentos que apresentam o número de tratamentos (I) igual a um quadrado perfeito ($I = k^2$). Exemplo disso é $I = k^2 = 5^2 = 25$, em que os tratamentos são distribuídos em repetições, nas quais cada repetição apresenta k blocos de k parcelas.

Seja, a seguir, o plano 10.1, mencionado em Cochran e Cox (1957), para o caso de um látice quadrado balanceado 3×3 planejado com doze blocos.

Plano 10.1. Látice quadrado balanceado com 3×3.

Blocos	Rep1			Blocos	Rep2			Blocos	Rep3			Blocos	Rep4		
B1	A	B	C	B4	A	D	G	B7	A	E	I	B10	A	F	H
B2	D	E	F	B5	B	E	H	B8	B	F	G	B11	B	D	I
B3	G	H	I	B6	C	F	I	B9	C	D	H	B12	C	E	G

No plano 10.1, observa-se que os tratamentos, de qualquer bloco da primeira repetição, se distribuem por todos os blocos da segunda e das demais repetições. Por exemplo, o terceiro bloco da primeira repetição que inclui os tratamentos G, H e I. Na segunda repetição, por sua vez, o tratamento G está no quarto bloco, o H no quinto e o I no sexto. Na terceira repetição, o G está no bloco 8, o H no nove e o I no sete. Na quarta repetição, o tratamento G está no bloco 12, o H no dez

e o I no 11. Esse tipo de distribuição ocorre em qualquer bloco de qualquer das quatro repetições, ou seja, os tratamentos de um bloco qualquer, de uma repetição qualquer, distribuem-se pelos blocos de qualquer outra repetição.

Esse tipo de distribuição dos tratamentos é o princípio fundamental do delineamento em látice (ou reticulado) quadrado. Repetições com essa propriedade são denominadas ortogonais. No caso do delineamento (plano 10.1) ocorrem quatro $(k+1)$ repetições ortogonais.

$$\lambda = \frac{J(k-1)}{I-1} = \frac{4(3-1)}{9-1} = 1$$

Tabela 10.2.1. ANOVA do delineamento em látice quadrado balanceado 3×3 .

F.V.	G.L.
Repetições	$k = 3$
Tratamentos (Não Aj.)	$k^2 - 1 = 8$
Blocos dentro das repetições (Aj.)	$k^2 - 1 = 8$
Resíduo	$(k-1)(k^2 - 1) = 16$
Total	$k^3 + k^2 - 1 = 35$

Considerar o plano do látice quadrado, $I = 16 = 4^2$, apresentado por Pimentel Gomes (2000), com três repetições (plano 10.5), apresentando doze blocos, constituídos cada um de quatro parcelas, mostrado a seguir.

Plano 10.5. Látice quadrado com $I = 4^2$ tratamentos com três repetições.

Blocos	Rep1				Blocos	Rep2				Blocos	Rep3			
B1	A	B	C	D	B5	A	E	I	M	B9	A	F	K	P
B2	E	F	G	H	B6	B	F	J	N	B10	B	E	L	O
B3	I	J	K	L	B7	C	G	K	O	B11	C	H	I	N
B4	M	N	O	P	B8	D	H	L	P	B12	D	G	J	M

$$\lambda = \frac{J(k-1)}{I-1} = \frac{3(4-1)}{16-1} = \frac{9}{15}$$

Tabela 10.2.2. ANOVA do delineamento látice quadrado ($I = 4^2$) com três repetições.

F.V.	G.L.
Repetições	$J - 1 = 2$
Tratamentos (Não Aj.)	$k^2 - 1 = 15$
Blocos dentro das repetições (Aj.)	$J(k-1) = 9$
Resíduo	$(k-1)(Jk - k - 1) = 21$
Total	$Jk^2 - 1 = 47$

Analisando o plano 10.5, verifica-se que os tratamentos alocados no bloco1, da repetição1, ou seja, A, B, C e D, estão distribuídos um para cada bloco da segunda repetição, assim como da terceira. Nota-se que esta mesma distribuição também ocorre com os tratamentos dos blocos 2, 3 e 4 de

qualquer repetição. Essa propriedade de distribuição dos tratamentos é denominada de ortogonal. No plano 10.5, ocorrem três repetições ortogonais.

No exemplo apresentado (plano 10.5), ocorre:

$$I = k^2 = 4^2 \quad b = mk = 12 \quad N = 48 = bk = mk^2$$

N : número total de parcelas;

b : número de blocos;

m : número de repetições ortogonais.

Quando $m = 2$, o reticulado é denominado simples ou duplo; se $m = 3$, é chamado de reticulado triplo.

De acordo com Pimentel Gomes e Garcia (2002), no látice $I = k^2$, quando $k \geq 2$, existem pelo menos três repetições ortogonais. Entretanto, no caso de k ser um número primo (tal como $k = 2, 3, 5, 7, 11, 13$) ou uma potência de um número primo, isto é, $k = 4 = 2^2$, $k = 8 = 2^3$ ou $k = 9 = 3^2$, pode-se ter sempre $k + 1$ repetições ortogonais. Para $k = 6$ ou $k = 12$ não há possibilidade de obter $k + 1$ repetições ortogonais.

Os reticulados quadrados $I = k^2$ com $(k + 1)$ repetições ortogonais são denominados látices balanceados ou equilibrados. Em outras palavras, qualquer tratamento ocorre juntamente com qualquer outro no experimento uma única vez. No plano 10.1, o delineamento látice apresentado é balanceado, enquanto, no plano 10.5, o apresentado é parcialmente balanceado.

Por outro lado, quando o número m de repetições ortogonais é inferior a $k + 1$, há pares de tratamentos denominados “primeiros associados”, que ocorrem no mesmo bloco uma única vez. Há outros chamados “segundos associados”, os quais são pares de tratamentos que não ocorrem em nenhum bloco.

No plano 10.5, os tratamentos C e D, que aparecem juntos no primeiro bloco, são primeiros associados, assim como os tratamentos G e O, que aparecem juntos no sétimo bloco. Entretanto, os tratamentos A e G não aparecem juntos em nenhum bloco do experimento. Por isso são denominados segundos associados.

O látice equilibrado ou balanceado planejado com $m = k + 1$ pode ser analisado como um delineamento em blocos incompletos equilibrados ou balanceados do tipo I, em que os blocos se agrupam em repetições.

De acordo com Pimentel Gomes e Garcia (2002), os delineamentos látices retangulares, cúbicos e os quadrados látices são raramente utilizados, enquanto o delineamento látice quadrado é o mais utilizado. Nesta obra, apenas será apresentado o delineamento látice quadrado ou reticulado quadrado, mas, para uma busca mais aprofundada sobre o tema, é recomendável consultar Pimentel Gomes (2000) e Cochran e Cox (1957).

De acordo com Pimentel Gomes e Garcia (2002), quando o número de parcelas (k) é um número primo ou uma potência de um número primo, pode-se obter até $m = k + 1$ repetições ortogonais, resultando em um reticulado quadrado equilibrado, que é um delineamento em blocos incompleto equilibrado. Por exemplo, se $k = 6$, é possível apenas produzir o reticulado duplo e o triplo.

Um reticulado apresentando um número m de repetições ortogonais pode ser repetido n vezes, resultando em:

$$I = k^2 \quad b = nmk \quad N = bk = nmk^2$$

Os reticulados, igualmente aos ensaios de blocos incompletos, podem ser analisados pelos métodos de análises intrablocos e com a recuperação da informação interblocos.

10.2.1. Análise intrablocos de reticulados quadrados

a) Soma de quadrados total (SQ_{TOTAL}), soma de quadrados das repetições (SQ_{REP}) e soma de quadrados de blocos dentro das repetições [$SQ_{BLOCOS \text{ D. REP (Não Aj.)}$].

$$SQ_{TOTAL} = \sum \sum x_{ij}^2 - C \quad SQ_{REP} = \frac{1}{k^2} \sum R^2 - C$$

$$SQ_{BLOCOS \text{ D. REP (Não Aj.)}} = \frac{1}{k} \sum B^2 - \frac{1}{k^2} R^2$$

B: Totais para os blocos;

R: Totais para repetições.

b) Soma de quadrados para tratamentos ajustados

O primeiro passo é obter os valores de $Q_i = kT_i - A_i$, calculando-se os totais de tratamentos (T_i) e as correções (A_i) que são as somas dos totais dos blocos que ocorrem no i -ésimo tratamento. O cálculo dos valores de $S_1(Q_i)$, $S_2(Q_i)$, ..., $S_m(Q_i)$ é obtido posteriormente, tal que:

$S_1(Q_i)$: é a soma dos valores de Q_i , concernentes a todos k tratamentos, os quais aparecem no mesmo bloco da primeira repetição, em que ocorre o i -ésimo tratamento;

$S_2(Q_i)$: é a soma dos valores de Q_i , concernentes a todos k tratamentos, os quais aparecem no mesmo bloco da segunda repetição, em que ocorre o i -ésimo tratamento;

$S_m(Q_i)$: é a soma dos valores de Q_i , concernentes a todos k tratamentos, os quais aparecem no mesmo bloco da m -ésima repetição, em que ocorre o i -ésimo tratamento.

Contudo, considerando $j = 1, 2, \dots, m$ e $j' = 1, 2, \dots, k$, define-se $S_{jj'}(Q)$ como a soma dos valores de Q_i para todos os tratamentos, os quais ocorrem no bloco J' , da repetição j , em que, se o i -ésimo tratamento aparecer nesse bloco, pode-se concluir que:

$$S_j(Q_i) = S_{jj'}(Q)$$

Seja o exemplo mostrado no plano 10.5:

Plano 10.5. Látice quadrada com $I = 4^2$ tratamentos com três repetições.

Blocos	Rep1				Blocos	Rep2				Blocos	Rep3			
B1	A	B	C	D	B5	A	E	I	M	B9	A	F	K	P
B2	E	F	G	H	B6	B	F	J	N	B10	B	E	L	O
B3	I	J	K	L	B7	C	G	K	O	B11	C	H	I	N
B4	M	N	O	P	B8	D	H	L	P	B12	D	G	J	M

Na repetição 1, considerando o bloco 1:

$$S_1(Q_A) = S_1(Q_B) = S_1(Q_C) = S_1(Q_D) = Q_A + Q_B + Q_C + Q_D = S_{11}(Q)$$

Na repetição 1, em relação ao bloco 2:

$$S_1(Q_E) = S_1(Q_F) = S_1(Q_G) = S_1(Q_H) = Q_E + Q_F + Q_G + Q_H = S_{12}(Q)$$

Assim por diante e igualmente para:

Na repetição 2, considerando o bloco 5:

$$S_2(Q_A) = S_2(Q_E) = S_2(Q_I) = S_2(Q_M) = Q_A + Q_E + Q_I + Q_M = S_{25}(Q)$$

Na repetição 2, considerando o bloco 6:

$$S_2(Q_B) = S_2(Q_F) = S_2(Q_J) = S_2(Q_N) = Q_B + Q_F + Q_J + Q_N = S_{26}(Q)$$

E assim sucessivamente.

c) Estimativa dos efeitos ajustados de tratamentos ($\hat{t}_i$):

$$\hat{t}_i = \frac{1}{mk} Q_i + \frac{1}{m(m-1)k^2} [S_1(Q_i) + S_2(Q_i) + \dots + S_m(Q_i)]$$

d) Estimativas das médias ajustadas de tratamentos (m'_i):

$$\hat{m}_i = \hat{m} + \hat{t}_i = \frac{G}{mk^2} + \hat{t}_i$$

e) Soma de quadrados de tratamentos ajustados:

$$SQ_{TRAT(Aj.)} = \frac{1}{k} \sum_{i=1}^I \hat{t}_i Q_i$$

f) Quadro da análise de variância

Tabela 10.2.1.1. ANOVA do látice quadrado ($I = 4^2$) com três repetições.

F.V.	G.L.	S.Q.	Q.M.
Repetições	$m - 1 = 2$	SQ_{REP}	
Blocos d. repetições	$m(k - 1) = 9$	$SQ_{BLOCOS D. REP(Não Aj.)}$	
Tratamentos (Aj.)	$k^2 - 1 = 15$	$SQ_{TRAT(Aj.)}$	$V_t = \frac{SQ_{TRAT(Aj.)}}{(k^2 - 1)}$
Resíduo	$(k - 1)(mk - k - 1) = 21$	Diferença	$\hat{\sigma}^2 = \frac{SQ_{RESÍDUO}}{(k - 1)(mk - k - 1)}$
Total	$mk^2 - 1 = 47$	SQ_{TOTAL}	

$F = \frac{V_t}{\hat{\sigma}^2}$: para comparar os contrastes entre tratamentos.

g) Teste de comparações múltiplas de Tukey

$$\Delta = q \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}_i - \hat{m}_u)}$$

De acordo com Pimentel Gomes (2000):

g1) Comparações entre pares de médias que ocorrem no mesmo bloco, os quais são denominados de pares primeiros associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left(1 + \frac{1}{k}\right)$$

g2) Comparações entre pares de médias que não ocorrem no mesmo bloco, os quais são denominados de pares segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)k}\right]$$

g3) Comparações entre pares de médias (pares primeiros ou segundos associados):

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)}\right]$$

h) Eficiência da análise intrablocos (EF) em relação aos blocos completos

É importante avaliar a eficiência do delineamento látice utilizado. Para obtê-la, é necessário efetuar a análise, considerando o Delineamento com Casualização em Blocos (DCB). A eficiência é dada pela fórmula:

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\sigma_{Ef}^2} \quad \sigma_{Ef}^2 : \text{Variância efetiva}$$

Pimentel Gomes e Garcia (2002) apresentam as três variâncias efetivas para obter a eficiência da análise intrablocos:

h1) Variância efetiva para pares primeiros associados:

$$\sigma_{Ef}^2 = \sigma^2 \left(1 + \frac{1}{k}\right)$$

h2) Variância efetiva para pares segundos associados:

$$\sigma_{Ef}^2 = \sigma^2 \left[1 + \frac{m}{(m-1)k}\right]$$

h3) Variância efetiva (média), indistintamente, usada para os pares primeiros e segundos associados:

$$\sigma_{Ef}^2 = \sigma^2 \left[1 + \frac{m}{(m-1)(k+1)}\right]$$

Exercício 10.2.1.1. Para a análise do látice quadrado (*Square Lattice*) seja o ensaio látice balanceado 3×3 (plano 10.1), para o estudo de competição de nove espécies florestais (dados simulados) com quatro repetições. A tabela 10.2.1.2 apresenta os valores do volume de madeira (variável resposta), em m^3 por parcela, medido na data da rotação economicamente aconselhável (doze anos). Efetuar a análise intrablocos.

Tabela 10.2.1.2. Dados do ensaio de competição de $I = 3^2 = 9$ espécies florestais.

Blocos	Rep1			Total		Rep2			Total
B1	E1=79,6	E2=66,6	E3=84,9	231,1	B4	E1=79,6	E4=55,7	E7=49,0	184,3
B2	E4=74,2	E5=30,7	E6=67,3	172,2	B5	E2=68,2	E5=58,9	E8=65,8	192,9
B3	E7=46,0	E8=65,8	E9=63,7	175,5	B6	E3=69,0	E6=61,3	E9=49,2	179,5
Blocos	Rep3					Rep4			
B7	E1=64,5	E5=35,0	E9=33,5	133,0	B10	E1=67,3	E6=59,1	E8=50,5	176,9
B8	E2=53,1	E6=65,2	E7=54,3	172,6	B11	E2=48,2	E4=56,1	E9=60,0	164,3
B9	E3=56,6	E4=47,4	E8=34,1	138,1	B12	E3=57,4	E5=39,2	E7=31,4	128,0

a) Soma de quadrados total, soma de quadrados dos blocos dentro das repetições e soma de quadrados para as repetições.

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 123.289,48 - \frac{2.048,4^2}{36} = 6.735,5200$$

$$SQ_{REP} = \frac{1}{k^2} (R_1^2 + R_2^2 + R_3^2 + R_4^2) - C = \frac{1}{9} (578,8^2 + 556,7^2 + 433,7^2 + 469,2^2) - C = 1.439,6689$$

$$SQ_{BLOCOS(Não\ Aj.)} = \frac{1}{3} (231,1^2 + 172,2^2 + \dots + 164,3^2 + 128,0^2) - \frac{2.048,4^2}{36}$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{3} \times 358.480,92 - \frac{2.048,4^2}{36} = 2.939,68$$

$$SQ_{BLOCOS D. REP (N\tilde{a}o Aj.)} = SQ_{BLOCOS (N\tilde{a}o Aj.)} - SQ_{REP} = 2.939,68 - 1.439,6689 = 1.500,0111$$

b) Soma de quadrados para tratamentos ajustados ($SQ_{TRAT(Aj.)}$)

b1) Obtenção dos totais dos blocos, em que ocorre o correspondente i -ésimo tratamento

Esp.	A_i	Total	Esp.	A_i	Total
E1	231,1+184,3+133,0+176,9	725,3	E6	172,2+179,5+172,6+176,9	701,2
E2	231,1+192,9+172,6+164,3	760,9	E7	175,5+184,3+172,6+128,0	660,4
E3	231,1+179,5+138,1+128,0	676,7	E8	175,5+192,9+138,1+176,9	683,4
E4	172,2+184,3+138,1+164,3	658,9	E9	175,5+179,5+133,0+164,3	652,3
E5	172,2+192,9+133,0+128,0	626,1			

b2) Obtenção dos valores de $Q_i = kT_i - A_i$, em que T_i ($i = 1, \dots, I$) representa o total dos tratamentos

Os valores calculados dos totais dos blocos — nos quais ocorre o i -ésimo tratamento (A_i) — e os totais para cada tratamento (T_i) para obter $Q_i = kT_i - A_i$, assim como os valores calculados de Q_i , constam na tabela 10.2.1.3.

b3) Cálculo da soma dos valores de Q_i , considerando todos os k tratamentos alocados no mesmo bloco da j -ésima repetição, ou seja, $S_1(Q_i)$, $S_2(Q_i)$, $S_3(Q_i)$ e $S_4(Q_i)$, em que, por exemplo, $S_1(Q_i)$ é a soma dos valores de Q_i para todos os k tratamentos, os quais aparecem no mesmo bloco da primeira repetição. No caso do bloco 1, da primeira repetição, segundo o plano 10.1, ocorrem os tratamentos E1, E2, E3. De acordo com a tabela 10.2.1.3:

$$S_1(Q_i) = 147,7 - 52,6 + 127,0 = 222,1$$

No caso do bloco 7 da terceira repetição, consoante o plano 10.1, ocorrem os tratamentos E1, E5, E9. Logo:

$$S_3(Q_i) = 147,7 - 134,7 - 33,1 = -20,1$$

No caso do bloco 10 da quarta repetição, conforme o plano 10.1, ocorrem os tratamentos E1, E6, E8. Logo:

$$S_3(Q_i) = 147,7 + 57,5 - 34,8 = 170,4$$

E assim procedendo:

Na repetição 1:

$$S_1(Q_1) = S_1(Q_2) = S_1(Q_3) = Q_1 + Q_2 + Q_3 = 147,7 - 52,6 + 127,0 = 222,1$$

$$S_1(Q_4) = S_1(Q_5) = S_1(Q_6) = Q_4 + Q_5 + Q_6 = 41,3 - 134,7 + 57,5 = -35,9$$

$$S_1(Q_7) = S_1(Q_8) = S_1(Q_9) = Q_7 + Q_8 + Q_9 = -118,3 - 34,8 - 33,1 = -186,2$$

Na repetição 2:

$$S_2(Q_1) = S_2(Q_4) = S_2(Q_7) = Q_1 + Q_4 + Q_7 = 147,7 + 41,3 - 118,3 = 70,7$$

$$S_2(Q_2) = S_2(Q_5) = S_2(Q_8) = Q_2 + Q_5 + Q_8 = -52,6 - 134,7 - 34,8 = -222,1$$

$$S_2(Q_3) = S_2(Q_6) = S_2(Q_9) = Q_3 + Q_6 + Q_9 = 127,0 + 57,5 - 33,1 = 151,4$$

Na repetição 3:

$$S_3(Q_1) = S_3(Q_5) = S_3(Q_9) = Q_1 + Q_5 + Q_9 = 147,7 - 134,7 - 33,1 = -20,1$$

$$S_3(Q_2) = S_3(Q_6) = S_3(Q_7) = Q_2 + Q_6 + Q_7 = -52,6 + 57,5 - 118,3 = -113,4$$

$$S_3(Q_3) = S_3(Q_4) = S_3(Q_8) = Q_3 + Q_4 + Q_8 = 127,0 + 41,3 - 34,8 = 133,5$$

Na repetição 4:

$$S_4(Q_1) = S_4(Q_6) = S_4(Q_8) = Q_1 + Q_6 + Q_8 = 147,7 + 57,5 - 34,8 = 170,4$$

$$S_4(Q_2) = S_4(Q_4) = S_4(Q_9) = Q_2 + Q_4 + Q_9 = -52,6 + 41,3 - 33,1 = -44,4$$

$$S_4(Q_3) = S_4(Q_5) = S_4(Q_7) = Q_3 + Q_5 + Q_7 = 127,0 - 134,7 - 118,3 = -126,0$$

Os valores calculados $S_1(Q_i)$, $S_2(Q_i)$, $S_3(Q_i)$ e $S_4(Q_i)$ constam na tabela 10.2.1.3.

b4) Estimativa dos efeitos ajustados de tratamentos ($\hat{t}_i$)

$$\hat{t}_i = \frac{1}{mk} Q_i + \frac{1}{m(m-1)k^2} [S_1(Q_i) + S_2(Q_i) + \dots + S_m(Q_i)]$$

$$\hat{t}_i = \frac{1}{4 \times 3} Q_i + \frac{1}{4(4-1) \times 3^2} [S_1(Q_i) + S_2(Q_i) + S_3(Q_i) + S_4(Q_i)]$$

$$\hat{t}_i = \frac{1}{12} Q_i + \frac{1}{108} [S_1(Q_i) + S_2(Q_i) + S_3(Q_i) + S_4(Q_i)]$$

Os valores dos efeitos ajustados de tratamentos ($\hat{t}_i$) constam na tabela 10.2.1.3.

Considerar a soma de quadrados de tratamentos ajustados a partir dos cálculos realizados:

$$SQ_{TRAT(Aj.)} = \frac{1}{k} \sum_{i=1}^I \hat{t}_i Q_i = \frac{1}{3} (8.907,6243) = 2.969,2081$$

c) Estimativas das médias ajustadas de tratamentos (m'_i)

$$\hat{m}_i = \hat{m} + \hat{t}_i = \frac{G}{mk^2} + \hat{t}_i \quad \hat{m} = \frac{G}{mk^2} = \frac{2.048,4}{4 \times 3^2} = 56,9 \quad \hat{m}_i = 56,9 + \hat{t}_i$$

Os valores calculados das médias ajustadas de tratamentos (m'_i) constam na tabela 10.2.1.3.

Tabela 10.2.1.3. Cálculos necessários para a análise intrablocos.

Esp.	T_i	A_i	$Q_i = KT_i - A_i$	$S_1(Q_i)$	$S_2(Q_i)$	$S_3(Q_i)$	$S_4(Q_i)$	$\hat{t}_i$	$\hat{m}_i$
E1	291,0	725,3	147,7	222,1	70,7	-20,1	170,4	16,4111	73,3111
E2	236,1	760,9	-52,6	222,1	-222,1	-113,4	-44,4	-5,8444	51,0556
E3	267,9	676,7	127,0	222,1	151,4	133,5	-126,0	14,1111	71,0111
E4	233,4	658,9	41,3	-35,9	70,7	133,5	-44,4	4,5889	61,4889
E5	163,8	626,1	-134,7	-35,9	-222,1	-20,1	-126,0	-14,9667	41,9333
E6	252,9	701,2	57,5	-35,9	151,4	-113,4	170,4	6,3889	63,2889
E7	180,7	660,4	-118,3	-186,2	70,7	-113,4	-126,0	-13,1444	43,7556
E8	216,2	683,4	-34,8	-186,2	-222,1	133,5	170,4	-3,8667	53,0333
E9	206,4	652,3	-33,1	-186,2	151,4	-20,1	-44,4	-3,6778	53,2222
Total	2048,4	6145,2	0	0	0	0	0	0	

d) Análise de variância

Tabela 10.2.1.4. ANOVA da análise intrablocos.

F.V.	G.L.	S.Q.	Q.M.	F
Repetições	3	1.439,6689	479,8896	9,29
Blocos d. repetições (Não Aj.)	8	1.500,0111	187,5014	3,63
Tratamentos (Aj.)	8	2.969,2082	371,1510	7,18**
Resíduo	16	826,6318	51,6645	
Total	35	6.735,5200		

Tabela 10.2.1.5. ANOVA realizada pelo SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Rep.	3	1439,66889	479,889630	9,29	0,0009
Blocos d. Rep. (Não Aj.)	8	1500,01111	187,501389	3,63	0,0135
Trat. (Aj.)	8	2969,20815	371,151019	7,18	0,0004
Blocos d. Rep. (Aj.) *	8	1159,44926	$\hat{V}_b = 144,931157$	2,81	0,0377
Resíduo	16	826,63185	$\hat{V}_r = \hat{\sigma}^2 = 51,66449$		
Total	35	6735,52000			

$$F_{8;16;0,01} = 3,89$$

A tabela 10.2.1.4 apresenta a ANOVA obtida dos cálculos realizados. A tabela 10.2.1.5 mostra a ANOVA pelo programa SAS, **Prog29**, em anexo no apêndice B. O programa SAS apresenta os valores referentes aos blocos dentro das repetições ajustados (*), os quais serão utilizados na análise com recuperação da informação interblocos.

e) Teste de comparação múltipla de Tukey

e1) Comparações entre pares de médias que ocorrem no mesmo bloco, os quais são denominados de pares primeiros associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left(1 + \frac{1}{k}\right) \quad \Delta = q_{9; 16; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)}$$

$$\Delta = q \sqrt{\frac{1}{2} \times \frac{2\hat{\sigma}^2}{m} \left(1 + \frac{1}{k}\right)} = 5,03 \sqrt{\frac{51,6645}{4} \left(1 + \frac{1}{3}\right)} = 20,8739$$

e2) Comparações entre pares de médias que não ocorrem no mesmo bloco, os quais são denominados de pares segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)k}\right] \quad \Delta = q_{9; 16; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)}$$

$$\Delta = q \sqrt{\frac{1}{2} \times \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)k}\right]} = 5,03 \sqrt{\frac{51,6645}{4} \left[1 + \frac{4}{(4-1)3}\right]} = 21,7262$$

e3) Comparações entre pares de médias, considerando, indistintamente, pares primeiros e segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)}\right] \quad \Delta = q_{9; 16; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)}$$

$$\Delta = q \sqrt{\frac{1}{2} \times \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)}\right]} = 5,03 \sqrt{\frac{51,6645}{4} \left[1 + \frac{4}{(4-1)(3+1)}\right]} = 20,8739$$

Tratando-se de um plano látice quadrado balanceado ou equilibrado, as comparações foram feitas, considerando os pares de médias que ocorrem no mesmo bloco (pares primeiros associados). Os valores da diferença mínima de significância (Δ) para os pares primeiros associados e para comparações entre pares de médias, em quaisquer casos (pares primeiros ou segundos associados) possuem o mesmo valor ($\Delta = 20,8739$).

Tratamentos	Médias ajustadas
E1	73,31 a
E3	71,01 ab
E6	63,29 abc
E4	61,49 abcd
E9	53,22 abcd
E8	53,03 abcd
E2	51,06 bcd
E7	43,76 cd
E5	41,93 d

Médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem pelo teste de Tukey. As espécies E1, E3, E6, E4, E9 e E8 apresentaram resultados não significantes. Contudo, é recomendável utilizar a espécie E1.

f) Eficiência, considerando a análise intrablocos

É importante avaliar a eficiência do delineamento látice utilizado. Para obtê-la, é necessário efetuar a análise, considerando o Delineamento Casualizado em Blocos (DCB). A eficiência é dada pela fórmula:

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\sigma_{Ef}^2} \quad \sigma_{Ef}^2 : \text{Variância efetiva}$$

Análise, considerando o experimento casualizado em blocos:

Tabela 10.2.1.6. Dados do experimento casualizado em blocos completos.

Rep.	E1	E2	E3	E4	E5	E6	E7	E8	E9
R1	79,6	66,6	84,9	74,2	30,7	67,3	46,0	65,8	63,7
R2	79,6	68,2	69,0	55,7	58,9	61,3	49,0	65,8	49,2
R3	64,5	53,1	56,6	47,4	35,0	65,2	54,3	34,1	33,5
R4	67,3	48,2	57,4	56,1	39,2	59,1	31,4	50,5	60,0

Tabela 10.2.1.7. ANOVA como em blocos completos.

F. V.	G.L.	S.Q.	Q.M.	F
Blocos = Rep.	3	1.439,6689	479,8896	5,80**
Trat.	8	3.309,7700	413,7213	5,00**
Resíduo	24	1.986,0811	82,7534	
Total	35	6.735,5200		

$$F_{8; 24; 0,01} = 3,36 \quad F_{3; 24; 0,01} = 4,72$$

A soma de quadrados do erro para o Delineamento Casualizado em Blocos (DCB) é igual à soma das somas de quadrados do erro intrablocos e à de blocos dentro de repetições ajustadas do delineamento em látice, quer dizer:

$$SQ_{RESÍDUO(DCB)} = SQ_{ERRO(INTRABLOCOS)} + SQ_{BLOCOS D.REP(Aj.)}$$

$$SQ_{RESÍDUO(DCB)} = 826,6319 + 1.159,4492 = 1.986,0811$$

f1) Estimativa da variância efetiva e eficiência para pares primeiros associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left(1 + \frac{1}{k}\right) = 51,6645 \left(1 + \frac{1}{3}\right) = 68,8860$$

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{82,7534}{68,8860} = 1,2013 \quad 120,13\%$$

f2) Estimativa da variância efetiva e eficiência para pares segundos associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)k} \right] = 51,6645 \left[1 + \frac{4}{(4-1)3} \right] = 74,6265$$

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{82,7534}{74,6265} = 1,1089 \text{ (110,89 \%)}$$

f3) Estimativa da variância efetiva e eficiência para pares primeiros ou segundos associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)(k+1)} \right] = 51,6645 \left[1 + \frac{4}{(4-1)(3+1)} \right] = 68,8860$$

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{82,7534}{68,8860} = 1,2013 \text{ (120,13 \%)}$$

De acordo com Pimentel Gomes e Garcia (1991), para que o látice tenha eficiência elevada, é necessário que o efeito dos blocos apresente grande magnitude. No caso de áreas experimentais uniformes ou de blocos mal localizados, o látice pode ter eficiência da ordem de 100%, sendo equivalente a um ensaio em blocos completos. Nesse ensaio, considerando ser um experimento balanceado (somente pares de tratamentos primeiros associados), conclui-se que a eficiência da análise intrablocos seja igual a 120,13 %.

Exercício 10.2.1.2. Para a análise do látice quadrado (*Square Lattice*), de acordo com o delineamento apresentado no plano 10.5, seja considerado um experimento com dados simulados com três repetições de competição de dezesseis ($I = 4^2$) espécies florestais. A tabela 10.2.1.8 apresenta os valores do volume de madeira (variável resposta) em m³ por parcela de 0,25 ha. As medições foram realizadas na data da rotação economicamente aconselhável (doze anos). Efetuar a análise intrablocos.

Tabela 10.2.1.8. Ensaio de competição de $I = 4^2$ espécies florestais.

Blocos	Repetição1				Total
B1	E1=31,9	E2=38,5	E3=35,4	E4=59,0	164,8
B2	E5=42,5	E6=38,2	E7=21,1	E8=27,6	129,4
B3	E9=24,6	E10=48,8	E11=23,7	E12=32,0	129,1
B4	E13=34,8	E14=51,4	E15=37,3	E16=42,9	166,4
Blocos	Repetição2				Total
B5	E1=40,2	E5=42,0	E9=40,8	E13=23,9	146,9
B6	E2=53,7	E6=43,4	E10=39,7	E14=31,0	167,8
B7	E3=47,3	E7=31,3	E11=33,7	E15=36,9	149,2
B8	E4=60,5	E8=42,0	E12=31,7	E16=23,6	157,8
Blocos	Repetição3				Total
B9	E1=45,3	E6=43,8	E11=39,3	E16=46,9	175,3
B10	E2=28,2	E5=28,7	E12=37,0	E15=37,8	131,7
B11	E3=25,7	E8=40,4	E9=24,2	E14=36,3	126,6
B12	E4=66,5	E7=55,9	E10=49,9	E13=33,4	205,7

a) Soma de quadrados total, soma de quadrados para as repetições e soma de quadrados para blocos dentro das repetições não ajustados.

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 76.464,09 - \frac{1.850,7^2}{48} = 5.108,0381$$

$$SQ_{REP} = \frac{1}{k^2} (R_1^2 + R_2^2 + R_3^2) - C$$

$$SQ_{REP} = \frac{1}{16} (589,7^2 + 621,7^2 + 639,3^2) - \frac{1.850,7^2}{48} = 79,0400$$

$$SQ_{BLOCOS(Não Aj.)} = \frac{1}{4} (164,8^2 + 129,4^2 + \dots + 126,6^2 + 205,7^2) - \frac{1.850,7^2}{48}$$

$$SQ_{BLOCOS (Não Aj.)} = \frac{1}{4} (291.572,13) - \frac{1.850,7^2}{48} = 1.536,9806$$

$$SQ_{BLOCOS D. REP(Não Aj.)} = SQ_{BLOCOS (Não Aj.)} - SQ_{REP}$$

$$SQ_{BLOCOS D. REP(Não Aj.)} = 1.536,9806 - 79,04 = 1.457,9406$$

b) Cálculo da soma dos quadrados para tratamentos ajustados ($SQ_{TRAT(Aj.)}$).

b1) Obtenção dos totais dos blocos em que ocorre o correspondente i -ésimo tratamento.

Na tabela 10.2.1.9, estão apresentados os totais dos blocos, em que ocorre o correspondente i -ésimo tratamento.

b2) Obtenção dos valores de $Q_i = kT_i - A_i$, tal que T_i ($i = 1, \dots, I$) é o total dos tratamentos. Os totais dos tratamentos e os valores de Q_i constam na tabela 10.2.1.10.

Tabela 10.2.1.9. Totais dos blocos em que aparece o i -ésimo tratamento.

Esp.	A_i	Total	Esp.	A_i	Total
E1	164,8+146,9+175,3	487,0	E9	129,1+146,9+126,6	402,6
E2	164,8+167,8+131,7	464,3	E10	129,1+167,8+205,7	502,6
E3	164,8+149,2+126,6	440,6	E11	129,1+149,2+175,3	453,6
E4	164,8+157,8+205,7	528,3	E12	129,1+157,8+131,7	418,6
E5	129,4+146,9+131,7	408,0	E13	66,4+146,9+205,7	519,0
E6	129,4+167,8+175,3	472,5	E14	166,4+167,8+126,6	460,8
E7	129,4+149,2+205,7	484,3	E15	166,4+149,2+131,7	447,3
E8	129,4+157,8+126,6	413,8	E16	166,4+157,8+175,3	499,5

b3) Cálculo da soma dos valores de Q_i , considerando todos os k tratamentos, alocados no mesmo bloco da j -ésima repetição. Ou seja, considerando $S_1(Q_i)$, $S_2(Q_i)$ e $S_3(Q_i)$, por exemplo, $S_1(Q_i)$ é a soma dos valores de Q_i para todos os k tratamentos que aparecem no mesmo bloco da primeira

repetição. No caso do bloco 1 da primeira repetição, segundo o plano 10.5, ocorrem os tratamentos E1, E2, E3 e E4. Logo:

$$S_1(Q_i) = -17,4 + 17,3 - 7 + 215,7 = 208,6$$

No caso do bloco 5 da segunda repetição, consoante o plano 10.5, verifica-se a ocorrência dos tratamentos E1, E5, E9 e E13. Portanto:

$$S_2(Q_i) = -17,4 + 44,8 - 44,2 - 150,6 = -167,4$$

No caso do bloco 9, da terceira repetição, conforme o plano 10.5, ocorrem os tratamentos E1, E6, E11 e E16. Assim:

$$S_3(Q_i) = -17,4 + 29,1 - 66,8 - 45,9 = -101,0$$

E assim sucessivamente.

Os valores de $S_1(Q_i)$, $S_2(Q_i)$ e $S_3(Q_i)$ constam na tabela 10.2.1.10.

No caso da repetição1:

$$S_1(Q_1) = S_1(Q_2) = S_1(Q_3) = S_1(Q_4) = Q_1 + Q_2 + Q_3 + Q_4 = 208,6$$

$$S_1(Q_5) = S_1(Q_6) = S_1(Q_7) = S_1(Q_8) = Q_5 + Q_6 + Q_7 + Q_8 = 49,0$$

$$S_1(Q_9) = S_1(Q_{10}) = S_1(Q_{11}) = S_1(Q_{12}) = Q_9 + Q_{10} + Q_{11} + Q_{12} = -75,8$$

$$S_1(Q_{13}) = S_1(Q_{14}) = S_1(Q_{15}) = S_1(Q_{16}) = Q_{13} + Q_{14} + Q_{15} + Q_{16} = -181,8$$

No caso da repetição2:

$$S_2(Q_1) = S_2(Q_5) = S_2(Q_9) = S_2(Q_{13}) = Q_1 + Q_5 + Q_9 + Q_{13} = -167,4$$

$$S_2(Q_2) = S_2(Q_6) = S_2(Q_{10}) = S_2(Q_{14}) = Q_2 + Q_6 + Q_{10} + Q_{14} = 111,4$$

$$S_2(Q_3) = S_2(Q_7) = S_2(Q_{11}) = S_2(Q_{15}) = Q_3 + Q_7 + Q_{11} + Q_{15} = -124,2$$

$$S_2(Q_4) = S_2(Q_8) = S_2(Q_{12}) = S_2(Q_{16}) = Q_4 + Q_8 + Q_{12} + Q_{16} = 180,2$$

No caso da repetição3:

$$S_3(Q_1) = S_3(Q_6) = S_3(Q_{11}) = S_3(Q_{16}) = Q_1 + Q_6 + Q_{11} + Q_{16} = -101,0$$

$$S_3(Q_2) = S_3(Q_5) = S_3(Q_{12}) = S_3(Q_{15}) = Q_2 + Q_5 + Q_{12} + Q_{15} = 47,0$$

$$S_3(Q_3) = S_3(Q_8) = S_3(Q_9) = S_3(Q_{14}) = Q_3 + Q_8 + Q_9 + Q_{14} = -11,0$$

$$S_3(Q_4) = S_3(Q_7) = S_3(Q_{10}) = S_3(Q_{13}) = Q_4 + Q_7 + Q_{10} + Q_{13} = 65,0$$

c) Estimativa dos efeitos ajustados de tratamentos ($\hat{t}_i$):

$$\hat{t}_i = \frac{1}{mk} Q_i + \frac{1}{m(m-1)k^2} [S_1(Q_i) + S_2(Q_i) + \dots + S_m(Q_i)]$$

$$\hat{t}_i = \frac{1}{3 \times 4} Q_i + \frac{1}{3(3-1)4^2} [S_1(Q_i) + S_2(Q_i) + S_3(Q_i)]$$

$$\hat{t}_i = \frac{1}{12} Q_i + \frac{1}{96} [S_1(Q_i) + S_2(Q_i) + S_3(Q_i)]$$

d) Estimativas das médias ajustadas de tratamentos (m'_i):

$$\hat{m}_i = \hat{m} + \hat{t}_i = \frac{G}{mk^2} + \hat{t}_i$$

$$\hat{m}_i = \hat{m} + \hat{t}_i = \frac{G}{mk^2} + \hat{t}_i = \frac{1.850,7}{3 \times 4^2} + \hat{t}_i = \frac{1.850,7}{48} + \hat{t}_i = 38,5563 + \hat{t}_i$$

Os valores dos efeitos ajustados de tratamentos ($\hat{t}_i$) e das médias ajustadas de tratamentos (m'_i) constam na tabela 10.2.1.10.

e) Soma de quadrados de tratamentos ajustados ($SQ_{TRAT(Aj.)}$).

$$SQ_{TRAT(Aj.)} = \frac{1}{k} \sum_{i=1}^J \hat{t}_i Q_i = \frac{1}{4} (9.275,3332) = 2.318,8333$$

Tabela 10.2.1.10. Resultados dos processamentos para a análise intrablocos.

Esp.	T_i	A_i	$Q_i = KT_i - A_i$	$S_1(Q_i)$	$S_2(Q_i)$	$S_3(Q_i)$	$\hat{t}_i$	$\hat{m}_i$
E1	117,4	487,0	-17,4	208,6	-167,4	-101,0	-2,0729	36,48
E2	120,4	464,3	17,3	208,6	111,4	47,0	5,2646	43,82
E3	108,4	440,6	-7,0	208,6	-124,2	-11,0	0,1812	38,74
E4	186,0	528,3	215,7	208,6	180,2	65,0	22,7021	61,26
E5	113,2	408,0	44,8	49,0	-167,4	47,0	2,9896	41,55
E6	125,4	472,5	29,1	49,0	111,4	-101,0	3,0438	41,60
E7	108,3	484,3	-51,1	49,0	-124,2	65,0	-4,3646	34,19
E8	110,0	413,8	26,2	49,0	180,2	-11,0	4,4563	43,01
E9	89,6	402,6	-44,2	-75,8	-167,4	-11,0	-6,3313	32,23
E10	138,4	502,6	51,0	-75,8	111,4	65,0	5,2979	43,85
E11	96,7	453,6	-66,8	-75,8	-124,2	-101,0	-8,7021	29,85
E12	100,7	418,6	-15,8	-75,8	180,2	47,0	0,2604	38,82
E13	92,1	519,0	-150,6	-181,8	-167,4	65,0	-15,5104	23,05
E14	118,7	460,8	14,0	-181,8	111,4	-11,0	0,3188	38,88
E15	112,0	447,3	0,7	-181,8	-124,2	47,0	-2,6396	35,92
E16	113,4	499,5	-45,9	-181,8	180,2	-101,0	-4,8938	33,66
Total	1.850,7	7.402,8	0	0	0	0	0	

f) Quadro da análise de variância

Tabela 10.2.1.11. ANOVA da análise intrablocos.

F.V.	G.L.	S.Q.	Q.M.	F
Repetições	2	79,0400	39,5200	0,66 <i>ns</i>
Blocos d. repetições (Não Aj.)	9	1.457,9406	161,9934	2,72*
Tratamentos (Aj.)	15	2.318,8333	154,5889	2,59*
Resíduo	21	1.252,2242	59,6297	
Total	47	5.108,0381		

Tabela 10.2.1.12. ANOVA realizada pelo SAS.

F.V.	G.L.	S.Q.	Q.M.	F	Pr > F
Rep.	2	79,040000	39,520000	0,66	0,5259
Blocos d. Rep. (Não Aj.)	9	1457,940625	161,993403	2,72	0,0284
Trat. (Aj.)	15	2318,833333	154,5889	2,59	0,0224
Blocos d. Rep. (Aj.) *	9	1251,649167	$\hat{V}_b = 139,0721$	2,33	0,0528
Resíduo	21	1252,224167	$\hat{V}_r = \hat{\sigma}^2 = 59,6297$		
Total	47	5108,038125			

$$F_{15; 21; 0,05} = 2,18 \quad F_{2; 21; 0,05} = 3,47 \quad F_{9; 21; 0,05} = 2,37$$

O programa SAS, **Prog30**, em anexo no apêndice B, apresenta os valores referentes aos blocos dentro das repetições ajustados (*), os quais serão utilizados na análise com recuperação da informação interblocos.

g) Teste de comparação múltipla de Tukey

g1) Comparação entre pares de médias que ocorrem no mesmo bloco, os quais são denominados de pares primeiros associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left(1 + \frac{1}{k}\right) \quad \Delta = q_{16; 21; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)}$$

$$\Delta = q \sqrt{\frac{1}{2} \times \frac{2\hat{\sigma}^2}{m} \left(1 + \frac{1}{k}\right)} = 5,47 \sqrt{\frac{59,6297}{3} \left(1 + \frac{1}{4}\right)} = 27,2655$$

g2) Comparação entre pares de médias que não ocorrem no mesmo bloco, os quais são denominados de pares segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)k}\right] \quad \Delta = q_{16; 21; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)}$$

$$\Delta = q \sqrt{\frac{1}{2} \times \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)k}\right]} = 5,47 \sqrt{\frac{59,6297}{3} \left[1 + \frac{3}{(3-1)4}\right]} = 28,5963$$

g3) Comparação entre pares de médias, indistintamente, denominados de primeiros ou segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)} \right] \quad \Delta = q_{16; 21; 0,05} \times \sqrt{\frac{1}{2} \hat{V}(\hat{m}_i - \hat{m}_u)}$$

$$\Delta = q \sqrt{\frac{1}{2} \times \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)} \right]} = 5,47 \sqrt{\frac{59,6297}{3} \left[1 + \frac{3}{(3-1)(4+1)} \right]} = 27,8054$$

Considerando que o experimento é parcialmente balanceado, as comparações entre pares de médias devem ser realizadas, utilizando critério que indistintamente sejam aplicáveis para os pares primeiros ou segundos associados.

Tratamentos	Médias ajustadas
E4	61,26 a
E10	43,85 ab
E2	43,82 ab
E8	43,01 ab
E6	41,60 ab
E5	41,55 ab
E14	38,88 ab
E12	38,82 ab
E3	38,74 ab
E1	36,48 ab
E15	35,92 ab
E7	34,19 ab
E16	33,66 ab
E9	32,23 b
E11	29,85 b
E13	23,05 b

As médias seguidas pela mesma letra, ao nível de 5% de significância, não diferem pelo teste de Tukey. As espécies E4, E10, E2, E8, E6, E5, E14, E12, E3, E1, E15, E7 e E16 apresentaram resultados não significantes. As informações mostram que é recomendável utilizar a espécie E4.

h) Cálculo da eficiência, considerando a análise intrablocos

Análise, considerando o experimento casualizado em blocos completos:

Tabela 10.2.1.13. Dados, considerando o ensaio em blocos completos.

Rep.	E1	E2	E3	E4	E5	E6	E7	E8
R1	31,9	38,5	35,4	59,0	42,5	38,2	21,1	27,6
R2	40,2	53,7	47,3	60,5	42,0	43,4	31,3	42,0
R3	45,3	28,2	25,7	66,5	28,7	43,8	55,9	40,4
Rep.	E9	E10	E11	E12	E13	E14	E15	E16
R1	24,6	48,8	23,7	32,0	34,8	51,4	37,3	42,9
R2	40,8	39,7	33,7	31,7	23,9	31,0	36,9	23,6
R3	24,2	49,9	39,3	37,0	33,4	36,3	37,8	46,9

A eficiência é dada pela fórmula a seguir:

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\sigma_{Ef}^2} \quad \sigma_{Ef}^2 : \text{Variância efetiva}$$

Tabela 10.2.1.14. ANOVA como em blocos completos.

F. V.	G.L.	S.Q.	Q.M.	F
Blocos = Rep.	2	79,0400	39,5200	0,47
Tratamentos	15	2.525,1248	168,3417	2,02*
Resíduo	30	2.503,8734	83,4624	
Total	47	5.108,0381		

$$F_{15; 30; 0,05} = 2,01 \quad F_{2; 30; 0,05} = 3,32$$

A soma de quadrados do erro para o delineamento casualizado em blocos (DCB) é igual à soma das somas de quadrados do erro intrablocos e à de blocos dentro de repetições ajustadas do delineamento em látice, isto é:

$$SQ_{RESÍDUO(DCB)} = SQ_{ERRO(INTRABLOCOS)} + SQ_{BLOCOS D.REP(Aj.)}$$

$$SQ_{RESÍDUO(DCB)} = 1.252,2242 + 1.251,6492 = 2.503,8734$$

h1) Estimativa da variância efetiva e eficiência para pares primeiros associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left(1 + \frac{1}{k}\right) = 59,6297 \left(1 + \frac{1}{4}\right) = 74,5371$$

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{83,4624}{74,5371} = 1,1197$$

h2) Estimativa da variância efetiva e eficiência para pares segundos associados:

Para avaliar a eficiência do delineamento látice utilizado, primeiramente é necessário efetuar a análise, considerando o Delineamento com Casualização em Blocos (DCB).

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)k}\right] = 59,6297 \left[1 + \frac{3}{(3-1)4}\right] = 81,9908$$

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{83,4624}{81,9908} = 1,018$$

h3) Estimativa da variância efetiva (média) e eficiência para pares primeiros e segundos associados.

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)(k+1)}\right] = 59,6297 \left[1 + \frac{3}{(3-1)(4+1)}\right] = 77,5186$$

$$EF = \frac{QM_{RESIDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{83,4624}{77,5186} = 1,0767$$

Para que o látice tenha eficiência elevada, é necessário que a magnitude do efeito dos blocos seja grande. Pimentel Gomes e Garcia (1991) comentam que no caso de áreas experimentais uniformes ou de blocos mal localizados, o látice pode ter eficiência da ordem de 100% e seria equivalente a um ensaio em blocos completos. Nesse exemplo, como o experimento é parcialmente balanceado, a eficiência é 107,67 %.

10.2.2. Análise com recuperação da informação interblocos de látices quadrados.

Em todos os experimentos em blocos, completos ou incompletos, o modelo matemático relativo à parcela que recebeu o tratamento i no bloco j é:

$$Y_{ij} = m + t_i + b_j + e_{ij}$$

m : média geral;

t_i : efeito do tratamento i ;

b_j : efeito do bloco j ;

e_{ij} : erro aleatório com variância σ^2 .

Na análise com recuperação da informação interblocos, o efeito de blocos b_j é considerado aleatório com variância σ_b^2 e igual para todos os blocos.

$$SQ_{BLOCOS(Aj.)} = SQ_{BLOCOS(Não Aj.)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(Não Aj.)}$$

Há dois métodos para a recuperação da informação interblocos, o método de Cochran e Cox (1957) e o de Bose (1954). Ambos utilizam as seguintes estatísticas:

$$V_b = \frac{SQ_{BLOCO(Aj.)}}{m(k-1)} \quad V_r = \sigma^2 = QM_{RESIDUO}$$

No método de Cochran e Cox (1957), os ajustes a serem feitos, na análise de variância com recuperação da informação interblocos, utilizam a estatística μ :

$$\mu = \frac{V_b - V_r}{k(m-1)V_b}$$

De acordo com Pimentel Gomes e Garcia (1991), entende-se:

$$E(\hat{V}_b) = \sigma^2 + \frac{k(m-1)\sigma_b^2}{m} \quad E(\hat{V}_r) = \sigma^2$$

De acordo com Cochran & Cox (1957), como $\sigma_b^2 \geq 0$, é esperado que $V_b \geq V_r$. Entretanto, eventualmente, devido à variação atribuída ao acaso, pode ocorrer que $V_b \leq V_r$. Nesses casos, admitindo-se não existir efeito dos blocos, considera-se $\mu = 0$.

Desse modo, a análise do látice é feita como um experimento casualizado em blocos completos.

Por outro lado, sempre ocorrerá:

$$\mu < \frac{1}{k(m-1)}$$

No método de Bose (1954), os ajustes a serem feitos, na análise com recuperação da informação interblocos, utilizam a estatística (a), cuja estimativa é dada pela fórmula:

$$\hat{a} = \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + k\hat{\sigma}_b^2}$$

$$\text{Como } V_b = \sigma^2 + \frac{k(m-1)\sigma_b^2}{m} \rightarrow \sigma_b^2 = \frac{m(V_b - \sigma^2)}{k(m-1)}$$

Substituindo $V_r = \sigma^2$ e $\sigma_b^2 = \frac{m(V_b - \sigma^2)}{k(m-1)}$ em $a = \frac{\sigma^2}{\sigma^2 + k\sigma_b^2}$, resulta:

$$a = \frac{(m-1)V_r}{mV_b - V_r}$$

É esperado que $V_b \geq V_r$. Contudo, devido à variação do acaso, pode ocorrer que $V_b \leq V_r$. Nesse caso, considera-se $a = 1$, admitindo-se não existir efeito dos blocos. Consequentemente, a análise do látice é feita como um experimento casualizado em blocos completos. Por outro lado, para $V_b > V_r$, obtém-se sempre $a > 0$. Assim sendo, o valor de (a) satisfaz as desigualdades: $0 < a < 1$.

No caso de $a = 0$, nenhum ajuste é feito e retorna-se à análise intrablocos, sem nenhuma alteração. Todavia, se $a = 1$, a análise com recuperação se torna idêntica à de blocos completos. Entretanto, para $0 < a < 1$, a análise com recuperação da informação interblocos difere das duas anteriormente citadas.

Pimentel Gomes e Garcia (1991) sugerem os seguintes procedimentos:

- i) Para $a > 0,8$, analisar como em blocos completos;
- ii) Para $a < 0,2$, usar a análise intrablocos;
- iii) No caso de $0,2 < a < 0,8$, usar a análise com recuperação da informação interblocos.

O valor do fator (a) de Bose (1954) pode ser obtido a partir da estatística μ de Cochran e Cox (1957):

$$a = \frac{1 - k(m-1)\mu}{1 + k\mu}$$

Para a análise com recuperação da informação interblocos, supondo o delineamento apresentado no plano 10.1, os seguintes cálculos devem ser realizados:

a) Soma de quadrados dos tratamentos não ajustados:

$$SQ_{TRAT(Não Aj.)} = \frac{1}{m}(T_1^2 + T_2^2 + \dots + T_l^2) - C$$

b) Soma de quadrados para blocos dentro das repetições ajustados:

$$SQ_{BLOCOS D. REP(Aj.)} = SQ_{BLOCOS D. REP(Não Aj.)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(Não Aj.)}$$

c) Calcular ($\hat{V}_b = QM_{BLOCOS D. REP(Aj.)}$) o quadrado médio para blocos ajustados dentro das repetições:

$$\hat{V}_b = \frac{SQ_{BLOCOS D. REP(Aj.)}}{m(k-1)}$$

d) Estimar o valor de ($\hat{a}$):

$$\hat{a} = \frac{(m-1)\hat{\sigma}^2}{m\hat{V}_b - \hat{\sigma}^2}$$

e) Estimar para cada tratamento o valor de M_i :

$$M_i = Q_i + \hat{a}A_i - \hat{a} \frac{G}{k}$$

f) Obter os efeitos ajustados para tratamentos ($\hat{t}'_i$):

$$\hat{t}'_i = \frac{1}{mk} M_i + \frac{1}{mk^2} \times \frac{1-\hat{a}}{m-1+\hat{a}} [S_1(M_i) + S_2(M_i) \dots + S_m(M_i)]$$

g) Obter os valores das médias ajustadas dos tratamentos ($\hat{m}'_i$):

$$\hat{m}'_i = \hat{m} + \hat{t}'_i = \frac{G}{mk^2} + \hat{t}'_i$$

h) Obter a soma de quadrados dos tratamentos para o novo ajustamento pela fórmula:

$$SQ_{TRAT(Aj.)} = \frac{\sum_{i=1}^l \hat{t}'_i M_i}{k} \rightarrow QM_{TRAT(Aj.)} = \frac{SQ_{TRAT(Aj.)}}{k^2 - 1}$$

Logo: $F = \frac{QM_{TRAT(Aj.)}}{\hat{\sigma}^2}$ com $(k^2 - 1)$ e $(k-1)(mk - k - 1)$ graus de liberdade.

i) Realização dos testes de comparações múltiplas

i1) Em pares de tratamentos primeiros associados, isto é, quando os dois tratamentos aparecem juntos em um mesmo bloco:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left(1 + \frac{\hat{V}_b - \hat{\sigma}^2}{k\hat{V}_b}\right)$$

i2) Em pares de tratamentos segundos associados, isto é, quando os dois tratamentos não aparecem juntos em um mesmo bloco:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)k} \times \frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b}\right]$$

i3) Indistintamente, para os casos de pares de tratamentos primeiros ou segundos associados, pode-se usar uma variância média:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)} \times \frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b}\right]$$

j) Cálculo da eficiência da análise de recuperação da informação interblocos

É importante avaliar a eficiência do delineamento látice utilizado. Para obtê-la, é necessário efetuar a análise, considerando o Delineamento Casualizado em Blocos (DCB). A eficiência é dada pela fórmula:

$$EF = \frac{QM_{RESIDUO(DCB)}}{\sigma_{Ef}^2} \cdot \sigma_{Ef}^2 : \text{Variância efetiva}$$

j1) Variância efetiva para pares primeiros associados:

$$\sigma_{Ef}^2 = \sigma^2 \left[1 + \frac{1}{k} \left(\frac{V_b - \sigma^2}{V_b}\right)\right]$$

j2) Variância efetiva para pares segundos associados:

$$\sigma_{Ef}^2 = \sigma^2 \left[1 + \frac{m}{(m-1)k} \left(\frac{V_b - \sigma^2}{V_b}\right)\right]$$

j3) Variância efetiva (média) para os pares, indistintamente, se primeiros ou segundos associados:

$$\sigma_{Ef}^2 = \sigma^2 \left[1 + \frac{m}{(m-1)(k+1)} \left(\frac{V_b - \sigma^2}{V_b}\right)\right]$$

Exercício 10.2.2.1. Para a análise do látice quadrado (*Square Lattice*), de acordo com o delineamento látice balanceado 3×3 apresentado no plano 10.1, seja considerado um experimento de competição de nove espécies florestais com quatro repetições. A tabela 10.2.2.1 apresenta os valores do volume de madeira (variável resposta), em m³ por parcela, medida na rotação economicamente aconselhável

(doze anos). No exercício 10.2.1.1, tais dados já foram analisados pelo método intrabloco. No presente exercício, efetuar a análise com recuperação da informação interblocos.

Tabela 10.2.2.1. Dados do ensaio de competição de $I = 3^2 = 9$ espécies florestais.

Blocos	Rep1			Total		Rep2			Total
B1	E1=79,6	E2=66,6	E3=84,9	231,1	B4	E1=79,6	E4=55,7	E7=49,0	184,3
B2	E4=74,2	E5=30,7	E6=67,3	172,2	B5	E2=68,2	E5=58,9	E8=65,8	192,9
B3	E7=46,0	E8=65,8	E9=63,7	175,5	B6	E3=69,0	E6=61,3	E9=49,2	179,5
Blocos	Rep3					Rep4			
B7	E1=64,5	E5=35,0	E9=33,5	133,0	B10	E1=67,3	E6=59,1	E8=50,5	176,9
B8	E2=53,1	E6=65,2	E7=54,3	172,6	B11	E2=48,2	E4=56,1	E9=60,0	164,3
B9	E3=56,6	E4=47,4	E8=34,1	138,1	B12	E3=57,4	E5=39,2	E7=31,4	128,0

a) Soma de quadrados total, soma de quadrados para as repetições, soma de quadrados dos blocos dentro das repetições não ajustados e a soma de quadrados de tratamentos não ajustados.

$$SQ_{TOTAL} = 123.289,48 - \frac{2.048,4^2}{36} = 6.735,5200$$

$$SQ_{REP} = \frac{1}{k^2} (R_1^2 + R_2^2 + R_3^2 + R_4^2) - C = \frac{1}{9} (5788^2 + 5567^2 + 4337^2 + 4692^2) - \frac{2.048,4^2}{36} = 1.439,6689$$

$$SQ_{BLOCOS(Não Aj.)} = \frac{1}{3} (231,1^2 + 172,2^2 + \dots + 164,3^2 + 128,0^2) - \frac{2.048,4^2}{36}$$

$$SQ_{BLOCOS(Não Aj.)} = \frac{1}{3} 358.480,92 - \frac{2.048,4^2}{36} = 2.939,68$$

$$SQ_{BLOCOS D. REP(Não Aj.)} = SQ_{BLOCOS(Não Aj.)} - SQ_{REP} = 2.939,68 - 1.439,6689 = 1.500,0111$$

$$SQ_{TRAT(Não Aj.)} = \frac{1}{m} (T_{E1}^2 + T_{E2}^2 + \dots + T_{E8}^2 + T_{E9}^2) - C = \frac{479.454,92}{4} - \frac{2.048,4^2}{36} = 3.309,7700$$

b) Soma de quadrados para blocos dentro das repetições ajustados ($SQ_{BLOCOS D. REP(Aj.)}$).

Do exercício 10.2.1.1, tabela 10.2.1.4, entende-se:

$$SQ_{TRAT(Aj.)} = \frac{1}{k} \sum_{i=1}^I \hat{t}_i Q_i = \frac{1}{3} (8.907,6243) = 2.969,2082$$

Então:

$$SQ_{BLOCOS D. REP(Aj.)} = SQ_{BLOCOS D. REP(Não Aj.)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(Não Aj.)}$$

$$SQ_{BLOCOS D. REP(Aj.)} = 1.500,0111 + 2.969,2081 - 3.309,7700 = 1.159,4493$$

$$QM_{BLOCO\ D.\ REP\ (A_j)} = \hat{V}_b = \frac{SQ_{BLOCO\ D.\ REP\ (A_j)}}{m(k-1)} = \frac{1.159,4492}{4(3-1)} = 144,9312$$

c) Obter o valor de $(\hat{a})$:

Do exercício 10.2.1.1, tabela 10.2.1.4, ocorre: $\hat{\sigma}^2 = 51,664491$

$$\hat{a} = \frac{(m-1)\hat{\sigma}^2}{m\hat{V}_b - \hat{\sigma}^2} = \frac{(4-1) \times 51,664491}{4 \times 144,931157 - 51,664491} = \frac{154,993473}{528,060137} = 0,29351481$$

Como recomendado, no caso de $0,2 < a < 0,8$, deve-se realizar a análise com recuperação da informação interblocos.

d) Obter o valor de (M_i)

$$M_i = Q_i + \hat{a}A_i - \hat{a}\frac{G}{k}$$

$$M_i = Q_i + 0,29351481A_i - 0,29351481 \times \frac{2.048,4}{3}$$

$$M_i = Q_i + 0,29351481A_i - 200,41191227$$

Os valores de M_i constam na tabela 10.2.2.2.

e) Obter $S_1(M_i)$, $S_2(M_i)$, $S_3(M_i)$ e $S_4(M_i)$

Na repetição1:

$$S_1(M_1) = S_1(M_2) = S_1(M_3) = M_1 + M_2 + M_3 = 160,174 - 29,677 + 125,210 = 255,707$$

$$S_1(M_4) = S_1(M_5) = S_1(M_6) = M_4 + M_5 + M_6 = 34,285 - 151,342 + 62,901 = -54,156$$

$$S_1(M_7) = S_1(M_8) = S_1(M_9) = M_7 + M_8 + M_9 = -124,875 - 34,624 - 42,052 = -201,551$$

Na repetição2:

$$S_2(M_1) = S_2(M_4) = S_2(M_7) = M_1 + M_4 + M_7 = 160,174 + 34,285 - 124,875 = 69,584$$

$$S_2(M_2) = S_2(M_5) = S_2(M_8) = M_2 + M_5 + M_8 = -29,677 - 151,342 - 34,624 = -215,643$$

$$S_2(M_3) = S_2(M_6) = S_2(M_9) = M_3 + M_6 + M_9 = 125,210 + 62,901 - 42,052 = 146,059$$

Na repetição3

$$S_3(M_1) = S_3(M_5) = S_3(M_9) = M_1 + M_5 + M_9 = 160,174 - 151,342 - 42,052 = -33,22$$

$$S_3(M_2) = S_3(M_6) = S_3(M_7) = M_2 + M_6 + M_7 = -29,677 + 62,901 - 124,875 = -91,651$$

$$S_3(M_3) = S_3(M_4) = S_3(M_8) = M_3 + M_4 + M_8 = 125,210 + 34,285 - 34,624 = 124,871$$

Na repetição 4

$$\begin{aligned} S_4(M_1) &= S_4(M_6) = S_4(M_8) = M_1 + M_6 + M_8 = 160,174 + 62,901 - 34,624 = 188,451 \\ S_4(M_2) &= S_4(M_4) = S_4(M_9) = M_2 + M_4 + M_9 = -29,677 + 34,285 - 42,052 = -37,444 \\ S_4(M_3) &= S_4(M_5) = S_4(M_7) = M_3 + M_5 + M_7 = 125,210 - 151,342 - 124,875 = -151,007 \end{aligned}$$

Os valores de $S_1(M_i)$, $S_2(M_i)$, $S_3(M_i)$ e $S_4(M_i)$ constam na tabela 10.2.2.2.

f) Obter o valor de ($\hat{t}'_i$)

$$\begin{aligned} \hat{t}'_i &= \frac{M_i}{mk} + \frac{1}{mk^2} \times \frac{1 - \hat{a}}{m - 1 + \hat{a}} [S_1(M_i) + S_2(M_i) + \dots + S_m(M_i)] \\ \hat{t}'_i &= \frac{M_i}{4 \times 3} + \frac{1}{4 \times 3^2} \times \frac{1 - 0,29351481}{4 - 1 + 0,29351481} [S_1(M_i) + S_2(M_i) + S_3(M_i) + S_4(M_i)] \\ \hat{t}'_i &= \frac{M_i}{12} + 0,005958554840 \times [S_1(M_i) + S_2(M_i) + S_3(M_i) + S_4(M_i)] \end{aligned}$$

Os valores de $\hat{t}'_i$ constam na tabela 10.2.2.2.

g) Estimar as médias ajustadas de tratamentos ($\hat{m}'_i$)

$$\begin{aligned} \hat{m}'_i &= \hat{m} + \hat{t}'_i = \frac{G}{mk^2} + \hat{t}'_i \quad \hat{m} = \frac{G}{mk^2} = \frac{2048,4}{4 \times 3^2} = 56,9 \\ \hat{m}'_i &= 56,9 + \hat{t}'_i \end{aligned}$$

Os valores de $\hat{m}'_i$ constam na tabela 10.2.2.2.

Tabela 10.2.2.2. Cálculos necessários para a análise de recuperação interblocos.

Esp.	A_i	Q_i	M_i	$S_1(M_i)$	$S_2(M_i)$	$S_3(M_i)$	$S_4(M_i)$	$\hat{t}'_i$	$\hat{m}'_i$
E1	725,3	147,7	160,174	255,707	69,584	-33,220	188,451	16,2111	73,1111
E2	760,9	-52,6	-29,677	255,707	-215,643	-91,651	-37,444	-3,0036	53,8964
E3	676,7	127,0	125,210	255,707	146,059	124,871	-151,007	12,6724	69,5724
E4	658,9	41,3	34,285	-54,156	69,584	124,871	-37,444	3,4700	60,3700
E5	626,1	-134,7	-151,342	-54,156	-215,643	-33,220	-151,007	-15,3172	41,5828
E6	701,2	57,5	62,901	-54,156	146,059	-91,651	188,451	6,3661	63,2661
E7	660,4	-118,3	-124,875	-201,551	69,584	-91,651	-151,007	-12,6385	44,2615
E8	683,4	-34,8	-34,624	-201,551	-215,643	124,871	188,451	-3,5043	53,3957
E9	652,3	-33,1	-42,052	-201,551	146,059	-33,220	-37,443	-4,2560	52,6440
Total	6145,2	0	0	0	0	0	0	0	

h) Soma de quadrados para tratamentos ajustados ($SQ_{TRAT(Aj.)}$).

$$SQ_{TRAT(Aj.)} = \frac{\sum_{i=1}^I \hat{t}'_i M_i}{k} = \frac{8.988,52}{3} = 2.996,1733$$

$$QM_{TRAT(Aj.)} = \frac{SQ_{TRAT(Aj.)}}{k^2 - 1} = \frac{2.996,1733}{8} = 374,5217$$

Deve ser observado que $SQ_{TRAT(Aj.)} = \frac{1}{k} \sum_{i=1}^I \hat{t}_i Q_i = \frac{1}{3}(8.907,6243) = 2.969,2082$ é a soma de quadrados dos tratamentos ajustados na análise intrabloco deste experimento (exercício 10.2.1.1), em que:

$$QM_{TRAT(Aj.)} = \frac{SQ_{TRAT(Aj.)}}{k^2 - 1} = \frac{2.969,2081}{3^2 - 1} = 371,1510$$

Por outro lado, usando o mesmo banco de dados para a análise com recuperação da informação interblocos, observa-se que o $QM_{TRAT(Aj.)} = 374,5217$, superando o valor obtido pela análise intrabloco $QM_{TRAT(Aj.)} = 371,1510$. Nesse caso, aumenta o valor de F que passa para $F = 7,25$, em vez de $F = 7,18$ (intrabloco).

Logo, considerar $F = \frac{QM_{TRAT(Aj.)}}{\hat{\sigma}^2}$ com $(k^2 - 1) = 8$ e $(k - 1)(mk - k - 1) = 16$ graus de liberdade.

Na tabela 10.2.1.4, ocorre: $\hat{\sigma}^2 = 51,6645$.

$$F = \frac{374,5417}{51,6645} = 7,25^{**}$$

$$F_{8;16;0,05} = 2,59 \quad F_{8;16;0,01} = 3,89$$

Tabela 10.2.2.3. Resultados da análise de recuperação da informação interblocos no SAS.

Análise de Variância			
F.V.	G.L.	S.Q.	Q.M.
Repetições	3	1439,67	479,89
Blocos dentro de Repetições (Aj.)	8	1159,45	144,93
Componente B	8	1159,45	144,93
Tratamentos (Não Aj.)	8	3309,77	413,72
Resíduo Intra Blocos	16	826,63	51,6645
Resíduo (Casualizado em Blocos)	24	1986,08	82,7534
Total	35	6735,52	192,44

Programa SAS, **Prog31**, em anexo no apêndice B. É importante observar, no programa **Prog31**, que no procedimento (PROC GLM DATA=LATICEINTER1;), a entrada dos dados dos tratamentos (T), das repetições (R) e dos blocos (B) apresenta um formato (tabela 10.2.2.4) diferente da mostrada no plano original 10.1. O programa não calcula o $SQ_{BLOCOS \ D.REP(Não \ Aj.)}$.

Tabela 10.2.2.4. Dados formatados usados para a análise no SAS.

R	B	T	Prod.	R	B	T	Prod.	R	B	T	Prod.	R	B	T	Prod.
1	1	1	79.6	2	1	4	55.7	3	3	3	56.6	4	2	2	48.2
1	2	5	30.7	2	2	8	65.8	3	1	9	33.5	4	3	5	39.2
1	3	9	63.7	2	1	1	79.6	3	1	5	35.0	4	3	3	57.4
1	2	4	74.2	2	2	5	58.9	3	2	7	54.3	4	1	8	50.5
1	3	8	65.8	2	3	9	49.2	3	1	1	64.5	4	1	6	59.1
1	3	7	46.0	2	2	2	68.2	3	2	6	65.2	4	2	9	60.0
1	1	3	84.9	2	3	6	61.3	3	3	8	34.1	4	1	1	67.3
1	1	2	66.6	2	3	3	69.0	3	2	2	53.1	4	2	4	56.1
1	2	6	67.3	2	1	7	49.0	3	3	4	47.4	4	3	7	31.4

i) Variância para o contraste entre duas médias

Dado que:

$$SQ_{BLOCOS D. REP(A_j)} = 1.500,0111 + 2.969,2081 - 3.309,7700 = 1.159,4493$$

Logo:

$$QM_{BLOCOS D. REP(A_j)} = \hat{V}_b = \frac{SQ_{BLOCOS D. REP(A_j)}}{m(k-1)} = \frac{1.159,4493}{4(3-1)} = 144,9312$$

Por outro lado $E(\hat{V}_r) = \sigma^2$, então: $\sigma^2 = 51,6645$.

i1) Para pares de tratamentos primeiros associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\sigma^2}{m} \left(1 + \frac{\hat{V}_b - \sigma^2}{k\hat{V}_b} \right)$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2 \times 51,6645}{4} \left(1 + \frac{144,9312 - 51,6645}{3 \times 144,9312} \right)$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 31,3735$$

i2) Para pares de tratamentos segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\sigma^2}{m} \left[1 + \frac{m}{(m-1)k} \left(\frac{\hat{V}_b - \sigma^2}{\hat{V}_b} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2 \times 51,6645}{4} \left[1 + \frac{4}{(4-1)3} \left(\frac{144,9312 - 51,6645}{144,9312} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 33,2205$$

i3) Para pares de tratamentos indistintamente, se primeiros associados ou segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2 \times 51,6645}{4} \left[1 + \frac{4}{(4-1)(3+1)} \left(\frac{144,9312 - 51,6645}{144,9312} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 31,3735$$

j) Eficiência, considerando a análise de recuperação da informação interblocos

Para avaliar a eficiência do delineamento látice utilizado, primeiramente, é necessário efetuar a análise, considerando o Delineamento Casualizado em Blocos (DCB). A eficiência é dada pela fórmula:

$$EF = \frac{QM_{RESIDUO(DCB)}}{\sigma_{Ef}^2} \quad \sigma_{Ef}^2: \text{Variância efetiva}$$

Análise de variância, considerando resultados do exercício 10.2.1.1 como um ensaio casualizado em blocos:

$$QM_{RESIDUO} = 82,7534$$

j1) Estimativa da variância efetiva e eficiência para pares primeiros associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{1}{k} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 51,6645 \left[1 + \frac{1}{3} \left(\frac{144,9312 - 51,6645}{144,9312} \right) \right] = 62,7469$$

$$EF = \frac{QM_{RESIDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{82,7534}{62,7469} = 1,3188 \rightarrow 131,88 \%$$

j2) Estimativa da variância efetiva e eficiência para pares segundos associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)k} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 51,6645 \left[1 + \frac{4}{(4-1)3} \left(\frac{144,9312 - 51,6645}{144,9312} \right) \right]$$

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)k} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 66,4411$$

$$EF = \frac{QM_{RESIDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{82,7534}{66,4411} = 1,2455 \rightarrow 124,55 \%$$

j3) Estimativa da variância efetiva (média) e eficiência para pares de tratamentos, indistintamente, se primeiros associados ou segundos associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)(k+1)} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 51,6645 \left[1 + \frac{4}{(4-1)(3+1)} \left(\frac{144,9312 - 51,6645}{144,9312} \right) \right]$$

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)(k+1)} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 83,7748$$

$$EF = \frac{QM_{RESÍDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{82,7534}{62,7469} = 1,3188 \rightarrow 131,88 \%$$

Tabela 10.2.2.5. Variância e eficiência dos pares de tratamentos primeiros associados no SAS.

Variância (média) no mesmo bloco: $\hat{V}(\hat{m}_i - \hat{m}_u)$	31,3735
Eficiência Relativa do RCBD	131,88

Para que o látice tenha eficiência elevada, é necessário que a magnitude do efeito dos blocos seja de grande proporção. Neste experimento, considerando se tratar de um ensaio balanceado, no qual só existem pares de tratamentos primeiros associados, conclui-se que a eficiência da análise com recuperação interblocos foi igual a 131,88 %. Ressalta-se, que na análise intrablocos, a eficiência foi de 120,13 %.

k) Teste de comparação múltipla de Tukey

k1) Para pares de tratamentos primeiros associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 31,3735 \quad q_{9;16;0,05} = 5,03$$

$$\Delta = q_{9;16;0,05} \times \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}_i - \hat{m}_u)} \quad \Delta = 5,03 \sqrt{\left(\frac{1}{2}\right) 31,3735} = 19,9221$$

k2) Para pares de tratamentos segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 33,2205 \quad q_{9;16;0,05} = 5,03$$

$$\Delta = q_{9;16;0,05} \times \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}_i - \hat{m}_u)} \quad \Delta = 5,03 \sqrt{\left(\frac{1}{2}\right) 33,2205} = 20,5001$$

k3) Para pares de tratamentos indistintamente, se primeiros associados ou segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 31,3735 \quad q_{9;16;0,05} = 5,03$$

$$\Delta = q_{9;16;0,05} \times \sqrt{\left(\frac{1}{2}\right) \hat{V}(\hat{m}_i - \hat{m}_u)} \quad \Delta = 5,03 \sqrt{\left(\frac{1}{2}\right) 31,3735} = 19,9221$$

Considerando que o experimento é balanceado, em que somente existem pares de tratamentos primeiros associados, deverá ser considerado $\Delta = 19,9221$.

Tratamentos	Médias ajustadas
E1	73,11 a
E3	69,57 ab
E6	63,27 abc
E4	60,37 abcd
E2	53,90 abcd
E8	53,40 abcd
E9	52,64 bcd
E7	44,26 cd
E5	41,58 d

Médias seguidas pela mesma letra não diferem, pelo teste de Tukey, ao nível de 5% de significância. As espécies E1, E3, E6, E4, E2 e E8 apresentaram resultados não significantes, mas pelo valor apresentado é recomendável utilizar E1.

Exercício 10.2.2.2. Análise do látice quadrado (*Square Lattice*), de acordo com o delineamento apresentado e analisado, plano 10.5, no exercício 10.2.1.2, pelo método intrablocos, sendo um experimento de competição de dezesseis ($I = 4^2$) espécies florestais, com três repetições. A tabela 10.2.2.6 apresenta os valores do volume de madeira, em m^3 por parcelas, medida na data da rotação economicamente aconselhável (doze anos). Analisar pelo método com recuperação da informação interblocos.

Tabela 10.2.2.6. Dados do ensaio de competição de $I = 4^2$ espécies florestais.

Blocos	Repetição 1				Total
B1	E1=31,9	E2=38,5	E3=35,4	E4=59,0	164,8
B2	E5=42,5	E6=38,2	E7=21,1	E8=27,6	129,4
B3	E9=24,6	E10=48,8	E11=23,7	E12=32,0	129,1
B4	E13=34,8	E14=51,4	E15=37,3	E16=42,9	166,4
Blocos	Repetição 2				Total
B5	E1=40,2	E5=42,0	E9=40,8	E13=23,9	146,9
B6	E2=53,7	E6=43,4	E10=39,7	E14=31,0	167,8
B7	E3=47,3	E7=31,3	E11=33,7	E15=36,9	149,2
B8	E4=60,5	E8=42,0	E12=31,7	E16=23,6	157,8
Blocos	Repetição 3				Total
B9	E1=45,3	E6=43,8	E11=39,3	E16=46,9	175,3
B10	E2=28,2	E5=28,7	E12=37,0	E15=37,8	131,7
B11	E3=25,7	E8=40,4	E9=24,2	E14=36,3	126,6
B12	E4=66,5	E7=55,9	E10=49,9	E13=33,4	205,7

a) Soma de quadrados total, soma de quadrados para as repetições, soma de quadrados para blocos dentro das repetições não ajustados e soma de quadrados dos tratamentos não ajustados.

$$SQ_{TOTAL} = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 - \frac{(\sum_{i=1}^I \sum_{j=1}^J x_{ij})^2}{IJ} = 76.464,09 - \frac{1.850,7^2}{48} = 5.108,0381$$

$$SQ_{REP} = \frac{1}{k^2} (R_1^2 + R_2^2 + R_3^2) - C = \frac{1}{16} (589,7^2 + 621,7^2 + 639,3^2) - \frac{1.850,7^2}{48} = 79,0400$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{4} (164,8^2 + 129,4^2 + \dots + 126,6^2 + 205,7^2) - \frac{1.850,7^2}{48}$$

$$SQ_{BLOCOS (N\tilde{a}o Aj.)} = \frac{1}{4} (291.572,13) - \frac{1.850,7^2}{48} = 1.536,9806$$

$$SQ_{BLOCOS D. REP(N\tilde{a}o Aj.)} = SQ_{BLOCOS (N\tilde{a}o Aj.)} - SQ_{REP} = 1.536,9806 - 79,04 = 1.457,9406$$

$$SQ_{TRAT(N\tilde{a}o Aj.)} = \frac{1}{m} (T_{E1}^2 + T_{E2}^2 + \dots + T_{E15}^2 + T_{E16}^2) - C = \frac{221.643,53}{3} - \frac{1.850,7^2}{48} = 2.525,1248$$

b) Soma de quadrados para blocos dentro das repetições ajustados:

$$SQ_{BLOCOS D. REP(Aj.)} = SQ_{BLOCOS D. REP(N\tilde{a}o Aj.)} + SQ_{TRAT(Aj.)} - SQ_{TRAT(N\tilde{a}o Aj.)}$$

$$SQ_{BLOCOS D. REP(Aj.)} = 1.457,9406 + 2.318,8333 - 2.525,1248 = 1.251,6492$$

$$QM_{BLOCOS(Aj.)} = \hat{V}_b = \frac{QM_{BLOCOS(Aj.)}}{m(k-1)} = \frac{1.251,6492}{9} = 139,0721$$

No exercício 10.2.1.2, para a análise de intrablocos deste experimento, a soma de quadrados dos tratamentos ajustados apresentou o seguinte resultado:

$$SQ_{TRAT(Aj.)} = \frac{1}{k} \sum_{i=1}^I \hat{t}_i Q_i = \frac{1}{4} (9.275,3332) = 2.318,8333$$

Em que:

$$QM_{TRAT(Aj.)} = \frac{SQ_{TRAT(Aj.)}}{k^2 - 1} = \frac{2.318,8333}{4^2 - 1} = 154,5889$$

c) Obter o valor de ($\hat{a}$):

$$\hat{a} = \frac{(m-1)\hat{\sigma}^2}{m\hat{V}_b - \hat{\sigma}^2} = \frac{(3-1)59,6297}{3 \times 139,0721 - 59,6297} = \frac{119,2594}{357,5866} = 0,3335$$

d) Obter o valor de (M_i):

$$M_i = Q_i + \hat{a}A_i - \hat{a}\frac{G}{k} \quad M_i = Q_i + 0,3335A_i - 0,3335 \times \frac{1850,7}{4}$$

$$M_i = Q_i + 0,3335A_i - 154,3021$$

e) Obter o valor de ($\hat{t}'_i$):

$$\hat{t}'_i = \frac{M_i}{mk} + \frac{1}{mk^2} \times \frac{1 - \hat{a}}{m - 1 + \hat{a}} [S_1(M_i) + S_2(M_i) + \dots + S_m(M_i)]$$

$$\hat{t}'_i = \frac{M_i}{12} + \frac{1}{48} \times \frac{1 - 0,3335}{3 - 1 + 0,3335} [S_1(M_i) + S_2(M_i) + \dots + S_m(M_i)]$$

$$\hat{t}'_i = \frac{M_i}{12} + 0,0059505047 \times [S_1(M_i) + S_2(M_i) + \dots + S_m(M_i)]$$

f) Estimar as médias ajustadas de tratamentos ($\hat{m}'_i$)

$$\hat{m}'_i = \hat{m} + \hat{t}'_i = \frac{G}{mk^2} + \hat{t}'_i \quad \hat{m} = \frac{G}{mk^2} = \frac{1850,7}{48} = 38,55625$$

Tabela 10.2.2.7. Resultados dos cálculos para a análise de recuperação interblocos.

Trat.	T_i	A_i	Q_i	M_i	$S_1(M_i)$	$S_2(M_i)$	$S_3(M_i)$	$\hat{t}'_i$	$\hat{m}_i$
E1	117,4	487,0	-17,4	-9,29	231,78	-178,78	-80,36	-0,9370	37,61
E2	120,4	464,3	17,3	17,84	231,78	127,92	9,48	3,6835	42,24
E3	108,4	440,6	-7,0	-14,36	231,78	-132,51	-55,32	-0,9351	37,62
E4	186,0	528,3	215,7	237,59	231,78	183,37	126,2	23,0205	61,58
E5	113,2	408,0	44,8	26,57	24,96	-178,78	9,48	1,3553	39,91
E6	125,4	472,5	29,1	32,38	24,96	127,92	-80,36	3,1299	41,69
E7	108,3	484,3	-51,1	-43,89	24,96	-132,51	126,2	-3,5465	35,01
E8	110,0	413,8	26,2	9,90	24,96	183,37	-55,32	1,7355	40,29
E9	89,6	402,6	-44,2	-64,24	-100,25	-178,78	-55,32	-7,3429	31,21
E10	138,4	502,6	51,0	64,32	-100,25	127,92	126,2	6,2756	44,83
E11	96,7	453,6	-66,8	-69,83	-100,25	-132,51	-80,36	-7,6824	30,87
E12	100,7	418,6	-15,8	-30,50	-100,25	183,37	9,48	-1,9907	36,57
E13	92,1	519,0	-150,6	-131,82	-156,49	-178,78	126,2	-12,2291	26,33
E14	118,7	460,8	14,0	13,38	-156,49	127,92	-55,32	0,6158	39,17
E15	112,0	447,3	0,7	-4,43	-156,49	-132,51	9,48	-2,0324	36,52
E16	113,4	499,5	-45,9	-33,62	-156,49	183,37	-80,36	-3,1199	35,44
Total	1850,70	7402,80	0		0	0	0	0	

g) Soma de quadrados para tratamentos ajustados:

$$SQ_{TRAT(Aj)} = \frac{\sum_{i=1}^I \hat{t}'_i M_i}{k} = \frac{9.074,09}{4} = 2.268,5$$

$$QM_{TRAT(Aj)} = \frac{SQ_{TRAT(Aj)}}{k^2 - 1} = \frac{2.268,5}{15} = 151,2333$$

Sendo o $QM_{TRAT(Aj.)} = 151,233$ da análise com recuperação da informação interblocos menor que o valor obtido pela análise intrablocos ($QM_{TRAT(Aj.)} = 154,5889$), diminuiu o valor de F que passa para $F = 2,54$, em vez de $F = 2,59$ (intrablocos). Desse modo, então:

$F = \frac{QM_{TRAT(Aj.)}}{\hat{\sigma}^2}$ com $(k^2 - 1) = 15$ e $(k - 1)(mk - k - 1) = 21$ graus de liberdade. Na tabela 10.2.1.11, ocorre $\hat{\sigma}^2 = 59,6297$.

$$F = \frac{151,2333}{59,6297} = 2,54 * \text{ com (15 e 21) graus de liberdade, pois: } F_{15; 21; 0,05} = 2,18.$$

Tabela 10.2.2.8. Resultados da análise de recuperação interblocos no SAS.

Análise de Variância			
F.V.	G.L.	S.Q.	Q.M.
Repetições	2	79,0400	39,5200
Blocos dentro de Repetições (Aj.)	9	1251,65	139,07
Componente B	9	1251,65	139,07
Tratamentos (Não Aj.)	15	2525,12	168,34
Resíduo Intra Blocos	21	1252,22	59,6297
Resíduo (Casualizado em Blocos)	30	2503,87	83,4624
Total	47	5108,04	108,68

Dados processados pelo programa SAS, **Prog32**, disponível no apêndice B. É importante observar, no programa **Prog32**, que no procedimento (PROC GLM DATA=LATICEINTER2;), a entrada dos dados dos tratamentos (T), das repetições (R) e dos blocos (B), apresenta um formato diferente da mostrada no plano original 10.5. O programa não calcula o $QM_{TRAT(Aj.)}$.

Tabela 10.2.2.9. Dados (plano 10.5) formatados para análise no SAS.

R	B	T	Prod.	R	B	T	Prod.	R	B	T	Prod.
1	1	1	31.9	2	1	1	40.2	3	1	1	45.3
1	1	2	38.5	2	1	5	42.0	3	1	6	43.8
1	1	3	35.4	2	1	9	40.8	3	1	11	39.3
1	1	4	59.0	2	1	13	23.9	3	1	16	46.9
1	2	5	42.5	2	2	2	53.7	3	2	2	28.2
1	2	6	38.2	2	2	6	43.4	3	2	5	28.7
1	2	7	21.1	2	2	10	39.7	3	2	12	37.0

Tabela 10.2.2.9. Dados (plano 10.5) formatados para análise no SAS (continuação).

R	B	T	Prod.	R	B	T	Prod.	R	B	T	Prod.
1	2	8	27.6	2	2	14	31.0	3	2	15	37.8
1	3	9	24.6	2	3	3	47.3	3	3	3	25.7
1	3	10	48.8	2	3	7	31.3	3	3	8	40.4
1	3	11	23.7	2	3	11	33.7	3	3	9	24.2
1	3	12	32.0	2	3	15	36.9	3	3	14	36.3
1	4	13	34.8	2	4	4	60.5	3	4	4	66.5
1	4	14	51.4	2	4	8	42.0	3	4	7	55.9
1	4	15	37.3	2	4	12	31.7	3	4	10	49.9
1	4	16	42.9	2	4	16	23.6	3	4	13	33.4

h) Variância para o contraste entre duas médias

$$QM_{BLOCOS(Aj)} = \hat{V}_b = \frac{QM_{BLOCOS(Aj)}}{m(k-1)} = \frac{1.251,65}{9} = 139,0721 \text{ e } \hat{\sigma}^2 = 59,6297$$

h1) Para dois tratamentos primeiros associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left(1 + \frac{\hat{V}_b - \hat{\sigma}^2}{k\hat{V}_b} \right)$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2 \times 59,6297}{3} \left(1 + \frac{139,07 - 59,6297}{4 \times 139,07} \right)$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 45,4302$$

h2) Para dois tratamentos segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)k} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2 \times 59,6297}{3} \left[1 + \frac{3}{(3-1)4} \left(\frac{139,07 - 59,6297}{139,07} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 48,2687$$

h3) Para dois pares de tratamentos indistintamente, primeiros ou segundos associados:

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2\hat{\sigma}^2}{m} \left[1 + \frac{m}{(m-1)(k+1)} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = \frac{2 \times 59,6297}{3} \left[1 + \frac{3}{(3-1)(4+1)} \left(\frac{139,07 - 59,6297}{139,07} \right) \right]$$

$$\hat{V}(\hat{m}_i - \hat{m}_u) = 46,5656$$

i) Eficiência, considerando a análise de recuperação da informação interblocos

$$EF = \frac{QM_{RESIDUO(DCB)}}{\sigma_{Ef}^2} \quad \sigma_{Ef}^2 : \text{Variância efetiva}$$

$$QM_{BLOCOS(Aj)} = \hat{V}_b = \frac{QM_{BLOCOS(Aj)}}{m(k-1)} = \frac{1.251,65}{9} = 139,0721 \text{ e } \hat{\sigma}^2 = 59,6297$$

Resultado da análise de variância, conforme o exercício 10.2.1.2, tabela 10.2.1.14, considerando como um ensaio casualizado em blocos:

$$QM_{RESIDUO} = 83,4624$$

i1) Estimativa da variância efetiva e eficiência para pares primeiros associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{1}{k} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 59,6297 \left[1 + \frac{1}{4} \left(\frac{139,0721 - 59,6297}{139,0721} \right) \right] = 68,1453$$

$$EF = \frac{QM_{RESIDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{83,4624}{68,1453} = 1,2248$$

i2) Estimativa da variância efetiva e eficiência para pares segundos associados:

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)k} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 59,6297 \left[1 + \frac{3}{(3-1)4} \left(\frac{139,0721 - 59,6297}{139,0721} \right) \right]$$

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)k} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 72,4031$$

$$EF = \frac{QM_{RESIDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{83,4624}{72,4031} = 1,1527$$

i3) Estimativa da variância efetiva (média) e eficiência para os pares de médias, indistintamente, primeiros ou segundos associados.

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)(k+1)} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 59,6297 \left[1 + \frac{3}{(3-1)(4+1)} \left(\frac{139,0721 - 59,6297}{139,0721} \right) \right]$$

$$\hat{\sigma}_{Ef}^2 = \hat{\sigma}^2 \left[1 + \frac{m}{(m-1)(k+1)} \left(\frac{\hat{V}_b - \hat{\sigma}^2}{\hat{V}_b} \right) \right] = 69,8484$$

$$EF = \frac{QM_{RESIDUO(DCB)}}{\hat{\sigma}_{Ef}^2} = \frac{83,4624}{69,8484} = 1,1949$$

Para que o látice tenha eficiência elevada, é necessário que a magnitude do efeito dos blocos seja grande. Neste experimento, considerando se tratar de um ensaio parcialmente balanceado e existindo pares de tratamentos primeiros e segundos associados, conclui-se que a eficiência da análise com recuperação interblocos seja igual a 119,49 %, ressaltando que, na análise intrablocos, a eficiência foi de 107,67 %.

Tabela 10.2.2.10. Variância para pares de médias e eficiência no SAS.

Variância para dois tratamentos primeiros associados	45,4302
Variância para dois tratamentos segundos associados	48,2687
Variância média: primeiros ou segundos associados	46,5656
Eficiência Relativa para RCBD	119,49

j) Teste de comparação múltipla de Tukey

j1) Para os tratamentos primeiros associados: $\hat{V}(\hat{m}_i - \hat{m}_u) = 45,4301$

$$\Delta = q_{16; 21; 0,05} \times \sqrt{\left(\frac{1}{2}\right)\hat{V}(\hat{m}_i - \hat{m}_u)} \quad \Delta = 5,47 \sqrt{\left(\frac{1}{2}\right)45,4301} = 26,07$$

j2) Para os tratamentos segundos associados: $\hat{V}(\hat{m}_i - \hat{m}_u) = 48,2686$

$$\Delta = q_{16; 21; 0,05} \times \sqrt{\left(\frac{1}{2}\right)\hat{V}(\hat{m}_i - \hat{m}_u)} \quad \Delta = 5,47 \sqrt{\left(\frac{1}{2}\right)48,2686} = 26,87$$

j3) Para os tratamentos em qualquer caso: $\hat{V}(\hat{m}_i - \hat{m}_u) = 46,5655$

$$\Delta = q_{16; 21; 0,05} \times \sqrt{\left(\frac{1}{2}\right)\hat{V}(\hat{m}_i - \hat{m}_u)} \quad \Delta = 5,47 \sqrt{\left(\frac{1}{2}\right)46,5655} = 26,39$$

TRATAMENTOS	MÉDIAS AJUSTADAS
E4	61,58 a
E10	44,83 ab
E2	42,24 ab
E8	41,69 ab
E6	40,29 ab
E5	39,91 ab
E14	39,17 ab
E12	37,62 ab
E3	37,61 ab
E1	36,57 ab
E15	36,52 ab
E7	35,44 ab
E16	35,01 b

E9	31,21 b
E11	30,87 b
E13	26,33 b

Foi considerado o valor $\Delta = 26,39$, haja vista tratar-se de um experimento parcialmente balanceado: Médias seguidas pela mesma letra não diferem, pelo teste de Tukey, ao nível de 5% de significância. As espécies E4, E10, E2, E8, E6, E5, E14, E12, E3, E1, E15 e E7 apresentaram resultados não significantes. As informações mostram que é recomendável utilizar a espécie E4.

10. 2.3. Repetição do delineamento em látice

Um reticulado com repetições ortogonais pode ser repetido n vezes. No caso de o experimento apresentar m repetições ortogonais, ao todo, ocorrerão $n \times m$ repetições. A análise é realizada pelos métodos já descritos.

Pimentel Gomes e Garcia (1991) apresentam a análise intrablocos de um experimento com *Eucalyptus grandis*, em látice duplo 6×6 duplicado, assim como a análise com recuperação da informação interblocos para um látice duplo 5×5 duplicado, com dados extraídos do livro de Cochran e Cox (1957).

Pimentel Gomes e Garcia (1991) comentam que para os látices de 6×6 ou de 10×10 só existem três repetições ortogonais e para os de 12×12 apenas quatro. Nessas condições, como seria possível instalar um látice de 10×10 (100 combinações) de um experimento florestal com 6 repetições? A solução seria considerar as três repetições ortogonais dadas por Cochran e Cox (1957), mas utilizando cada uma delas duas vezes, originando a denominação de um látice triplo duplicado. Alternativamente, podem ser usadas duas repetições ortogonais, cada uma delas três vezes, obtendo-se assim um látice duplo (ou simples) triplicado.

Em linhas gerais, a análise dos dados segue os métodos já vistos anteriormente, em que as especificações para o uso dos mesmos já foram vistas. Isto é:

- a) Analisar como casualizados em blocos completos;
- b) Realizar a análise intrablocos;
- c) Realizar a análise com recuperação da informação interblocos.

Há a possibilidade de incluir, em cada bloco de um delineamento reticulado (ou látice), um tratamento testemunha, a ser casualizado juntamente com os demais tratamentos (chamados regulares) em cada bloco. A análise intrablocos do delineamento assim obtida foi publicada por Pimentel Gomes e Viégas (1978). Esses autores apresentaram a análise de um reticulado quadrado com 5×5 tratamentos, incluindo uma testemunha em cada bloco.

Oliveira e Barbin (1988) mostraram o método de análise intrablocos para o caso de ensaios, em reticulado quadrado, com adição de alguns tratamentos comuns. Oliveira (1990) apresenta os métodos de análise intrablocos e com recuperação da informação interblocos, para o caso de ensaios em reticulado quadrado, com adição de alguns tratamentos comuns a todos os blocos.

As expressões na análise de variância foram apresentadas para as várias somas de quadrados, as médias de tratamentos ajustadas para blocos, assim como a variância da estimativa de contraste entre as médias de dois tratamentos. Um exemplo numérico foi apresentado para ilustrar os métodos propostos.

10.3. BIBLIOGRAFIA

BOSE, R.C. **Least squares aspects of analysis of variance**. Raleigh: NCSU, 1954. 87 p.

COCHRAN, W. G.; COX, G. M. **Experimental designs**. 2ª ed. New York: John Wiley, 1957. 611 p.

MORAES, M. L. T.; VALERA, F. P.; MORI, E. S.; KAGEAMA, P. Y. Aspectos práticos dos delineamentos em látice. Série técnica - **IPEF** - Piracicaba v.5, n.14, p. 1-21. Fev. 1988.

OLIVEIRA, A.C.; BARBIN, D. Experimentos em reticulado quadrado com alguns tratamentos comuns adicionados em cada bloco: análise intrablocos. **Pesquisa Agropecuária Brasileira**, Brasília, **23 (7)**: 717 - 723, 1988.

OLIVEIRA, A.C. Experimentos em reticulado quadrado com alguns tratamentos comuns adicionados em cada bloco: análise com recuperação da informação interblocos. **Pesquisa agropecuária brasileira**, Brasília, 25 (3): 289 - 298, 1990.

PIMENTEL GOMES, F. **Curso de estatística experimental**. 14ª ed. São Paulo: Nobel, 2000. 475 p.

PIMENTEL GOMES, F.; GARCIA, C. H. Experimentos em látice: planejamento e análise por meio de “pacotes” estatísticos. Série técnica – **IPEF** - Piracicaba, 7(23): 1-69, 1991.

PIMENTEL GOMES, F.; GARCIA, C. H. **Estatística aplicada a experimentos agrônômicos e florestais**. Piracicaba: FEALQ. (Biblioteca de Ciências Agrárias Luiz de Queiroz - Volume 11), 2002. 309 p.

PIMENTEL GOMES, F.; VIÉGAS, G.P. Experiments in square lattice with a common treatment in all blocks. **Revista da Agricultura**, Piracicaba, 53: 34-43, 1978.



CAPÍTULO 11

METODOLOGIA DE SUPERFÍCIES DE RESPOSTA

Este capítulo apresenta definições, conceitos e aplicações sobre a metodologia de superfícies de resposta. Trata ainda sobre a análise de superfície de resposta de um ensaio 3×3 . Cálculo das doses economicamente aconselháveis. Apresenta, também, a análise de superfície de resposta de um ensaio $3 \times 3 \times 3$, o uso da análise canônica na otimização de modelos na forma quadrática, o delineamento composto central, o planejamento Box-Behnken e o exemplo de aplicação de um fatorial $(\frac{1}{5})5^3$.

11.1. ANÁLISE DE SUPERFÍCIE DE RESPOSTA DE UM ENSAIO FATORIAL 3×3

É usual denominar a representação geométrica das equações que caracterizam os dados de fenômenos naturais como curva de resposta. Seja considerado o caso de um experimento fatorial de adubação 3×3 com P e K em um ensaio inteiramente ao acaso.

Seja considerada a equação de regressão polinomial de segundo grau (segunda ordem), considerando $P = X_1$ e $K = X_2$:

$$Y_i = a_0 + a_1X_1 + a_2X_2 + a_{11}X_1^2 + a_{22}X_2^2 + a_{12}X_1X_2 + e_i \quad 11.1.1$$

É frequente chamar de superfície de resposta a representação geométrica destes tipos de modelos em estudos de adubação, em que X_1 se refere aos níveis de P e X_2 aos níveis de K .

11.1.1. Estimativas dos parâmetros da regressão

Considerando que os níveis de P e K são equidistantes e x_1 e x_2 são as variáveis transformadas, em que q_1 e q_2 são respectivamente as diferenças entre os níveis sucessivos de X_1 e X_2 :

$$x_1 = \frac{X_1 - \bar{X}_1}{q_1} \quad x_2 = \frac{X_2 - \bar{X}_2}{q_2}$$

De acordo com a expressão 11.1.1, resulta em:

$$Y_i = a_0 + a_1x_1 + a_2x_2 + a_{11}x_1^2 + a_{22}x_2^2 + a_{12}x_1x_2 + e_i \quad 11.1.1.1$$

Considerando que o fator P ocorre nos níveis $X_1 = 0, 20, 40$ e o fator K nos níveis $X_2 = 0, 30, 60$, em kg por parcela, assim como os valores da variável resposta também em kg por parcela, então:

$$\bar{X}_1 = 20 \text{ e } q_1 = 20 \quad \bar{X}_2 = 30 \text{ e } q_2 = 30$$

Consequentemente: $x_1 = (-1, 0, 1)$ e $x_2 = (-1, 0, 1)$.

Considerando:

	K_0	K_1	K_2
P_0	P_0K_0	P_0K_1	P_0K_2
P_1	P_1K_0	P_1K_1	P_1K_2
P_2	P_2K_0	P_2K_1	P_2K_2

Pode-se estruturar a matriz X do modelo da regressão 11.1.1.1 da seguinte forma:

Tabela 11.1.1.1. Matriz X do modelo da regressão 11.1.1.1.

Trat.	a_0	a_1	a_2	a_{11}	a_{22}	a_{12}
y_{00}	1	-1	-1	1	1	1
y_{01}	1	-1	0	1	0	0
y_{02}	1	-1	1	1	1	-1
y_{10}	1	0	-1	0	1	0
y_{11}	1	0	0	0	0	0
y_{12}	1	0	1	0	1	0
y_{20}	1	1	-1	1	1	-1
y_{21}	1	1	0	1	0	0
y_{22}	1	1	1	1	1	1

Seja considerado o modelo de regressão na forma matricial:

$$Y_v = X\beta_v + e_v$$

Considerando o modelo 11.1.1.1, tem-se:

$$Y_v = \begin{bmatrix} y_{00} \\ y_{01} \\ y_{02} \\ y_{10} \\ y_{11} \\ y_{12} \\ y_{20} \\ y_{21} \\ y_{22} \end{bmatrix} = \begin{bmatrix} P_0 K_0 \\ P_0 K_1 \\ P_0 K_2 \\ P_1 K_0 \\ P_1 K_1 \\ P_1 K_2 \\ P_2 K_0 \\ P_2 K_1 \\ P_2 K_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{11} \\ a_{22} \\ a_{12} \end{bmatrix} + \begin{bmatrix} e_{00} \\ e_{01} \\ e_{02} \\ e_{10} \\ e_{11} \\ e_{12} \\ e_{20} \\ e_{21} \\ e_{22} \end{bmatrix}$$

Como $X'X\hat{\beta}_v = X'Y_v$ é a solução de mínimos quadrados e J o número de repetições, tem-se:

$$X'X = J \times \begin{bmatrix} 9 & 0 & 0 & 6 & 6 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 6 & 0 & 0 & 6 & 4 & 0 \\ 6 & 0 & 0 & 4 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

$$X'Y_v = \begin{bmatrix} G \\ P_2 - P_0 \\ K_2 - K_0 \\ P_0 + P_2 \\ K_0 + K_2 \\ P_0K_0 + P_2K_2 - (P_0K_2 + P_2K_0) \end{bmatrix}$$

O vetor $Y'_v = [y_{00} \ y_{01} \ y_{02} \ y_{10} \ y_{11} \ y_{12} \ y_{20} \ y_{21} \ y_{22}]$ é formado pelos valores dos totais dos tratamentos.

$$X'X\hat{\beta}_v = J \times \begin{bmatrix} 9 & 0 & 0 & 6 & 6 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 6 & 0 & 0 & 6 & 4 & 0 \\ 6 & 0 & 0 & 4 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix} \times \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{12} \end{bmatrix} = X'Y_v = \begin{bmatrix} G \\ P_2 - P_0 \\ K_2 - K_0 \\ P_0 + P_2 \\ K_0 + K_2 \\ P_0K_0 + P_2K_2 - (P_0K_2 + P_2K_0) \end{bmatrix}$$

Para reparametrizar o modelo 11.1.1.1, e objetivando transformar $X'X$ em uma matriz diagonal, usa-se o valor “ c ” nos termos quadráticos, em que I é o número de combinações. Logo:

$$Y_i = a_0 + a_1x_1 + a_2x_2 + a_{11}(x_1^2 - c) + a_{22}(x_2^2 - c) + a_{12}x_1x_2 + e_i \quad 11.1.1.2$$

Da tabela 11.1.1.1, resulta a tabela 11.1.1.2:

Tabela 11.1.1.2. Matriz X do modelo reparametrizado.

Trat.	a_0	a_1	a_2	a_{11}	a_{22}	a_{12}
y_{00}	1	-1	-1	$1 - c$	$1 - c$	1
y_{01}	1	-1	0	$1 - c$	$0 - c$	0
y_{02}	1	-1	1	$1 - c$	$1 - c$	-1
y_{10}	1	0	-1	$0 - c$	$1 - c$	0
y_{11}	1	0	0	$0 - c$	$0 - c$	0
y_{12}	1	0	1	$0 - c$	$1 - c$	0
y_{20}	1	1	-1	$1 - c$	$1 - c$	-1
y_{21}	1	1	0	$1 - c$	$0 - c$	0
y_{22}	1	1	1	$1 - c$	$1 - c$	1

Somando os elementos da coluna a_{11} ou da a_{22} e igualando a zero, resulta em:

$$6(1 - c) + 3(0 - c) = 0 \quad 6 - 6c - 3c = 0 \quad c = \frac{6}{9} = \frac{2}{3}$$

Tabela 11.1.1.3. Matriz X da regressão para o modelo reparametrizado ($c = \frac{2}{3}$).

Trat.	a_0	a_1	a_2	a_{11}	a_{22}	a_{12}
y_{00}	1	-1	-1	$1 - c = \frac{1}{3}$	$1 - c = \frac{1}{3}$	1
y_{01}	1	-1	0	$1 - c = \frac{1}{3}$	$0 - c = -\frac{2}{3}$	0
y_{02}	1	-1	1	$1 - c = \frac{1}{3}$	$1 - c = \frac{1}{3}$	-1
y_{10}	1	0	-1	$0 - c = -\frac{2}{3}$	$1 - c = \frac{1}{3}$	0
y_{11}	1	0	0	$0 - c = -\frac{2}{3}$	$0 - c = -\frac{2}{3}$	0
y_{12}	1	0	1	$0 - c = -\frac{2}{3}$	$1 - c = \frac{1}{3}$	0
y_{20}	1	1	-1	$1 - c = \frac{1}{3}$	$1 - c = \frac{1}{3}$	-1
y_{21}	1	1	0	$1 - c = \frac{1}{3}$	$0 - c = -\frac{2}{3}$	0
y_{22}	1	1	1	$1 - c = \frac{1}{3}$	$1 - c = \frac{1}{3}$	1
Total	9	0	0	0	0	0

Portanto:

$$\begin{aligned}
 XX' &= J \times \begin{bmatrix} 9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix} \\
 XX'\hat{\beta}_v &= J \times \begin{bmatrix} 9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix} \times \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{12} \end{bmatrix} = XY_v = \begin{bmatrix} G \\ P_2 - P_0 \\ K_2 - K_0 \\ \frac{1}{3}(P_0 - 2P_1 + P_2) \\ \frac{1}{3}(K_0 - 2K_1 + K_2) \\ P_0K_0 + P_2K_2 - (P_0K_2 + P_2K_0) \end{bmatrix} \\
 \hat{\beta}_v &= \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{12} \end{bmatrix} = \frac{1}{J} \begin{bmatrix} G/9 \\ \frac{1}{6}(P_2 - P_0) \\ \frac{1}{6}(K_2 - K_0) \\ \frac{1}{6}(P_0 - 2P_1 + P_2) \\ \frac{1}{6}(K_0 - 2K_1 + K_2) \\ \frac{1}{4}[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)] \end{bmatrix}
 \end{aligned}$$

11.1.2. Análise de variância para o modelo 11.1.1.2

a) Soma de quadrados dos tratamentos

$$SQ_{TRAT} = \frac{Y_v'Y_v}{J} - \frac{G^2}{9J}$$

b) Soma de quadrados dos desvios da regressão (SQ_{DREG})

$$SQ_{DREG} = \frac{Y_v'Y_v}{J} - \frac{G^2}{9J} - \left\{ \frac{(P_2 - P_0)^2}{6J} + \frac{(K_2 - K_0)^2}{6J} + \frac{(P_0 - 2P_1 + P_2)^2}{18J} + \frac{(K_0 - 2K_1 + K_2)^2}{18J} + \frac{[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)]^2}{4J} \right\}$$

$$SQ_{DREG} = SQ_{TRAT} - [SQ_{P'} + SQ_{K'} + SQ_{P''} + SQ_{K''} + SQ_{(P' \times K')}]$$

Matriz de dispersão dos coeficientes:

$$\hat{D} = \begin{bmatrix} V(\hat{a}_0) & 0 & 0 & 0 & 0 & 0 \\ 0 & V(\hat{a}_1) & 0 & 0 & 0 & 0 \\ 0 & 0 & V(\hat{a}_2) & 0 & 0 & 0 \\ 0 & 0 & 0 & V(\hat{a}_{11}) & 0 & 0 \\ 0 & 0 & 0 & 0 & V(\hat{a}_{22}) & 0 \\ 0 & 0 & 0 & 0 & 0 & V(\hat{a}_{12}) \end{bmatrix}$$

Tabela 11.1.2.1. ANOVA da regressão.

F. V.	G.L.	S.Q.
Tratamentos	(8)	SQ_{TRAT}
P'	1	$\frac{(P_2 - P_0)^2}{6J}$
P''	1	$\frac{(P_0 - 2P_1 + P_2)^2}{18J}$
K'	1	$\frac{(K_2 - K_0)^2}{6J}$
K''	1	$\frac{(K_0 - 2K_1 + K_2)^2}{18J}$
$P' \times K'$	1	$\frac{[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)]^2}{4J}$
Desvios da regressão	3	$SQ_{TRAT} - [SQ_{P'} + SQ_{K'} + SQ_{P''} + SQ_{K''} + SQ_{(P' \times K')}]$
Resíduo	$9(J - 1)$	$SQ_{TOTAL} - SQ_{TRAT}$
Total	$9J - 1$	$\sum_{i=1}^I \sum_{j=1}^J y_{ij}^2 - \frac{G^2}{9J}$

P' e P'' : efeitos linear e quadrático de P ;

K' e K'' : efeitos linear e quadrático de K ;

$P' \times K'$: efeito da interação entre P' e K' .

11.1.3. Cálculo das doses economicamente aconselháveis

Dado o modelo de superfície de resposta:

$$Y_i = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^k a_{ii} x_i^2 + \sum_i \sum_j a_{ij} x_i x_j + e_i \quad i < j$$

Em que:

$\sum_{i=1}^k a_i x_i$: componentes de primeira ordem;

$\sum_{i=1}^k \sum_{i=1}^k a_{ii} x_i^2 + \sum_i \sum_j a_{ij} x_i x_j$: componentes de segunda ordem.

Seja considerada a equação estimada:

$$\hat{Y}_i = \hat{a}_0 + \hat{a}_1 x_1 + \hat{a}_2 x_2 + \hat{a}_{11} x_1^2 + \hat{a}_{22} x_2^2 + \hat{a}_{12} x_1 x_2$$

Considerando um ensaio de adubação 3^2 de uma espécie florestal. Seja considerada RL a equação da receita líquida:

$$RL = w\hat{Y}_i - cf - t_1 q_1 x_1 - t_2 q_2 x_2$$

$$\frac{RL}{w} = \hat{Y}_i - \frac{cf}{w} - \frac{t_1 q_1 x_1}{w} - \frac{t_2 q_2 x_2}{w}$$

w : preço da madeira ($m^3 \text{ ha}^{-1}$);

t_1 : preço do fósforo (P_2O_5) por ha;

t_2 : preço do potássio (K_2O) por ha;

cf : custos fixos.

Tomando a diferencial de primeira ordem da equação da receita líquida:

$$\frac{dRL}{w} = d\hat{Y}_i - \frac{t_1 q_1}{w} dx_1 - \frac{t_2 q_2}{w} dx_2$$

$$\hat{Y}_i = \hat{a}_0 + \hat{a}_1 x_1 + \hat{a}_2 x_2 + \hat{a}_{11} x_1^2 + \hat{a}_{22} x_2^2 + \hat{a}_{12} x_1 x_2$$

$$d\hat{Y}_i = \hat{a}_1 dx_1 + \hat{a}_2 dx_2 + 2\hat{a}_{11} x_1 dx_1 + 2\hat{a}_{22} x_2 dx_2 + \hat{a}_{12} x_2 dx_1 + \hat{a}_{12} x_1 dx_2$$

$$d\hat{Y}_i = (\hat{a}_1 + 2\hat{a}_{11} x_1 + \hat{a}_{12} x_2) dx_1 + (\hat{a}_2 + 2\hat{a}_{22} x_2 + \hat{a}_{12} x_1) dx_2$$

$$\frac{dRL}{w} = (\hat{a}_1 + 2\hat{a}_{11} x_1 + \hat{a}_{12} x_2) dx_1 + (\hat{a}_2 + 2\hat{a}_{22} x_2 + \hat{a}_{12} x_1) dx_2 - \frac{t_1 q_1}{w} dx_1 - \frac{t_2 q_2}{w} dx_2$$

$$\frac{dRL}{w} = (\hat{a}_1 + 2\hat{a}_{11}x_1 + \hat{a}_{12}x_2 - \frac{t_1q_1}{w})dx_1 + (\hat{a}_2 + 2\hat{a}_{22}x_2 + \hat{a}_{12}x_1 - \frac{t_2q_2}{w})dx_2$$

$$dx_1=0 \text{ e } dx_2=0 \rightarrow \begin{cases} \hat{a}_1 + 2\hat{a}_{11}x_1^* + \hat{a}_{12}x_2^* - \frac{t_1q_1}{w} = 0 \\ \hat{a}_2 + 2\hat{a}_{22}x_2^* + \hat{a}_{12}x_1^* - \frac{t_2q_2}{w} = 0 \end{cases}$$

Sistema de equações para obter as doses economicamente aconselháveis (x_1^*, x_2^*):

$$\begin{cases} 2\hat{a}_{11}x_1^* + \hat{a}_{12}x_2^* = \frac{t_1q_1}{w} - \hat{a}_1 \\ \hat{a}_{12}x_1^* + 2\hat{a}_{22}x_2^* = \frac{t_2q_2}{w} - \hat{a}_2 \end{cases}$$

Tomando a diferencial de segunda ordem para verificar que se trata de um ponto de máximo:

$$\frac{dRL}{w} = (\hat{a}_1 + 2\hat{a}_{11}x_1 + \hat{a}_{12}x_2 - \frac{t_1q_1}{w})dx_1 + (\hat{a}_2 + 2\hat{a}_{22}x_2 + \hat{a}_{12}x_1 - \frac{t_2q_2}{w})dx_2$$

$$\frac{d^2RL}{w} = 2\hat{a}_{11}d^2x_1 + 2\hat{a}_{22}d^2x_2 + \hat{a}_{12}dx_1dx_2 + \hat{a}_{12}dx_1dx_2$$

$$\frac{d^2RL}{w} = 2\hat{a}_{11}d^2x_1 + 2\hat{a}_{22}d^2x_2 + 2\hat{a}_{12}dx_1dx_2$$

$$\frac{d^2RL}{2w} = \hat{a}_{11}d^2x_1 + \hat{a}_{22}d^2x_2 + \hat{a}_{12}dx_1dx_2$$

$$\frac{d^2RL}{2w} = \begin{bmatrix} dx_1 & dx_2 \end{bmatrix} \begin{bmatrix} \hat{a}_{11} & \frac{\hat{a}_{12}}{2} \\ \frac{\hat{a}_{12}}{2} & \hat{a}_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix}$$

Em que A é a matriz Hessiana:

$$A = \begin{bmatrix} \hat{a}_{11} & \frac{\hat{a}_{12}}{2} \\ \frac{\hat{a}_{12}}{2} & \hat{a}_{22} \end{bmatrix}$$

- a) No caso da matriz A ser definida positiva, então (x_1^*, x_2^*) corresponde a um ponto de mínimo;
- b) Se a matriz A for definida negativa, então (x_1^*, x_2^*) corresponde a um ponto de máximo;
- c) Se a matriz A for não definida, então (x_1^*, x_2^*) não é máximo e nem mínimo, ou seja, é um ponto de sela (Saddle point).

Por outro lado, em alguns casos, pode ser obtido o ponto estacionário sem considerar a maximização da receita líquida.

11.1.4. Uso da análise canônica na otimização de modelos na forma quadrática

Na análise da forma de uma superfície de resposta e na localização das regiões de condições ótimas, muitas vezes, no estudo de otimização de experimentos com fatores quantitativos, é importante transformar, na análise de uma superfície de resposta, o modelo escrito na forma quadrática para a forma canônica.

Dado o modelo de superfície de resposta de segunda ordem escrito na forma quadrática:

$$Y_i = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^k a_{ii} x_i^2 + \sum_i \sum_j a_{ij} x_i x_j + e_i \quad i < j$$

O objetivo é transformar o modelo Y_i para a forma canônica:

$$Y_i = Y_0 + \lambda_1 u_1^2 + \lambda_2 u_2^2 + \cdots + \lambda_k u_k^2 + e_i$$

A equação na forma canônica não apresenta os termos lineares e nem os termos de interação. O modelo é escrito apenas com termos quadráticos, em que o objetivo é obter a natureza do ponto estacionário. Trata-se de uma transformação ortogonal dos eixos correspondentes das variáveis independentes $x_1, x_2, \dots, x_k$, cuja origem é $(0, 0, \dots, 0)$ para as variáveis $u_1, u_2, \dots, u_k$. Ou seja, há uma mudança do eixo original para um novo eixo de sistema de coordenadas, tal que a nova origem passa a ser exatamente o ponto estacionário:

$$x'_{v0} = [x_{01} \quad x_{02} \quad \cdots \quad x_{0k}]$$

Seja considerado o modelo escrito matricialmente da forma quadrática Y_i :

$$Y_i = a_0 + x'_v a_v + x'_v A x_v + e_i$$

a_0 : parâmetro algébrico;

$x'_v a_v$: termos de primeira ordem;

$x'_v A x_v$: termos de segunda ordem.

$$x_v = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad a_v = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12}/2 & \dots & a_{1k}/2 \\ a_{12}/2 & a_{22} & \dots & a_{2k}/2 \\ \dots & \dots & \dots & \dots \\ a_{1k}/2 & \dots & \dots & a_{kk} \end{bmatrix}$$

De $\frac{\partial Y_i}{\partial x_1} = \frac{\partial Y_i}{\partial x_2} = \dots = \frac{\partial Y_i}{\partial x_k} = 0$, existe o ponto estacionário: $x'_{v0} = [x_{01} \quad x_{02} \quad \dots \quad x_{0k}]$.

Em que:

$$\frac{\partial Y_i}{\partial x_v} = \frac{\partial}{\partial x_v} (a_0 + x'_v a_v + x'_v A x_v)$$

$$\frac{\partial Y_i}{\partial x_v} = \frac{\partial}{\partial x_v} (a_0 + x'_v a_v + x'_v A x_v) = a_v + 2A x_v$$

$$a_v + 2A x_{v0} = 0$$

$$x_{v0} = -\frac{1}{2} A^{-1} a_v \text{ é o ponto crítico ou estacionário.}$$

Na análise canônica, a superfície quadrática de segunda ordem $Y_i = a_0 + x'_v a_v + x'_v A x_v + e_i$ é transformada para a forma canônica.

$$Y_i = Y_0 + \lambda_1 u_1^2 + \lambda_2 u_2^2 + \dots + \lambda_k u_k^2 + e_i$$

Objetivando facilitar a determinação da natureza do ponto estacionário, o modelo é escrito apenas em função de termos quadráticos.

Quando a forma quadrática $Y_i = a_0 + x'_v a_v + x'_v A x_v + e_i$ é transformada para a forma canônica $Y_i = Y_0 + \lambda_1 u_1^2 + \lambda_2 u_2^2 + \dots + \lambda_k u_k^2 + e_i$, não existirão mais os termos lineares e nem os de interação.

O procedimento é considerar a translação da superfície de resposta da origem $x_1, x_2, \dots, x_k = (0, 0, \dots, 0)$ para a nova origem $x_{v0} = -\frac{1}{2} A^{-1} a_v$, ou seja, para o ponto estacionário. Daí, a função resposta é formulada a partir das novas variáveis $u_1, u_2, \dots, u_k$ e Y_0 , em que Y_0 é o valor da produção em x_{v0} . Os novos eixos correspondem aos eixos principais do sistema de contorno, em que os parâmetros ajustados são os autovalores $\lambda_1, \lambda_2, \dots, \lambda_k$. Seja considerado o procedimento para obter a redução de uma forma quadrática para uma canônica:

Da álgebra de matrizes, se $\lambda_1, \lambda_2, \dots, \lambda_k$ são as raízes características ou autovalores (todos reais) de uma matriz simétrica real B , então é possível efetuar uma transformação ortogonal $x_v = P u_v$, tal que a forma quadrática real é transformada para a forma canônica $\lambda_1 u_1^2 + \lambda_2 u_2^2 + \dots + \lambda_k u_k^2$. Isto é, a forma quadrática $Q = x'_v B x_v$ é transformada para uma forma com uma matriz diagonal, em que os seus elementos diagonais são as raízes características da matriz B .

Os sinais e as grandezas dos autovalores $\lambda_1, \lambda_2, \dots, \lambda_k$ ajudam a determinar a natureza dos pontos estacionários. Ao mesmo tempo, a relação entre $u_1, u_2, \dots, u_k$ e $x_1, x_2, \dots, x_k$ é importante para indicar regiões úteis para exploração.

Definindo uma nova variável $z_v = x_v - x_{v0}$ para efetuar a translação da superfície de resposta para uma nova origem x_{v0} , logo $x_v = z_v + x_{v0}$, em que $x'_v = [x_1 \ x_2 \ \dots \ x_k]$ e $x_{v0} = -0,5A^{-1}a_v$.

De $Y_i = a_0 + x'_v a_v + x'_v A x_v + e_i$, resulta:

$$Y_i = a_0 + (z'_v + x'_{v0})a_v + (z'_v + x'_{v0})A(z_v + x_{v0}) + e_i$$

$$Y_i = a_0 + z'_v a_v + x'_{v0} a_v + z'_v A z_v + x'_{v0} A z_v + z'_v A x_{v0} + x'_{v0} A x_{v0} + e_i$$

$$Y_i = a_0 + x'_{v0} a_v + x'_{v0} A x_{v0} + z'_v a_v + z'_v A x_{v0} + x'_{v0} A z_v + z'_v A z_v + e_i$$

Dado que: $Y_0 = a_0 + x'_{v0} a_v + x'_{v0} A x_{v0}$ e $z'_v A x_{v0} = x'_{v0} A z_v$, então:

$$Y_i = a_0 + x'_{v0} a_v + x'_{v0} A x_{v0} + z'_v a_v + z'_v A z_v + z'_v A x_{v0} + x'_{v0} A z_v + e_i$$

$$Y_i = Y_0 + z'_v a_v + z'_v A z_v + 2z'_v A x_{v0} + e_i$$

$$Y_i = Y_0 + z'_v (a_v + 2A x_{v0}) + z'_v A z_v + e_i$$

Dado que: $x_{v0} = -\frac{1}{2}A^{-1}a_v$, então:

$$Y_i = Y_0 + z'_v [a_v - 2A(A^{-1} \frac{a_v}{2})] + z'_v A z_v + e_i$$

$$Y_i = Y_0 + z'_v A z_v + e_i$$

Feita a transformação de x_v em z_v , o próximo passo será transformar z_v em u_v , o que permitirá obter a expressão que relaciona as variáveis x_v com as variáveis u_v .

Para obter a transformação de $z_v = [z_1, z_2, \dots, z_k]$ em $u_v = [u_1, u_2, \dots, u_k]$, é necessário obter a matriz $M = [m_{v1} \ m_{v2} \ \dots \ m_{vk}]$, tal que $u_v = M' z_v$. As colunas da matriz M são os autovetores associados às raízes características $\lambda_1, \lambda_2, \dots, \lambda_k$ da matriz A .

Logo, determina-se a matriz $M = [m_{v1} \ m_{v2} \ \dots \ m_{vk}]$, tal que $u_v = M' z_v$.

m_{vi} : são os autovetores normalizados associados aos autovalores λ_i .

$m'_{vi} = [m_{1i} \ m_{2i} \ \dots \ m_{ki}]$, tal que $m_{1i}^2 + m_{2i}^2 + \dots + m_{ki}^2 = m'_{vi} m_{vi} = 1$.

Como usar os autovalores estimados para classificar o tipo de ponto estacionário?

- a) Se $\hat{\lambda}_i < 0$, para $i = 1, 2, \dots, k$, significa que quando se caminha em qualquer direção, a partir do ponto estacionário, ocorre um decréscimo da resposta, isto é, o ponto estacionário é o ponto de resposta de máximo da superfície ajustada, com concavidade para baixo;
- b) Se $\hat{\lambda}_i > 0$, para $i = 1, 2, \dots, k$, significa que quando se caminha em qualquer direção, a partir do ponto estacionário, produz-se um acréscimo da resposta, ou seja, o ponto estacionário é o ponto de resposta de mínimo da superfície ajustada, com concavidade para cima;
- c) Se os autovalores estimados possuem sinais diferentes, o ponto estacionário não é nem máximo e nem mínimo.

Seja considerado o exemplo $\hat{Y} = \hat{Y}_0 + \hat{\lambda}_1 u_1^2 + \hat{\lambda}_2 u_2^2$:

Se $\hat{\lambda}_1 < 0$ e $\hat{\lambda}_2 > 0$ (ponto de sela), significa que um movimento no eixo u_1 , a partir de x_{v0} em qualquer direção, origina um decréscimo na resposta estimada, enquanto um movimento no eixo u_2 provoca um acréscimo na resposta estimada.

Se $\hat{\lambda}_1 < 0$ e $\hat{\lambda}_2 < 0$, em que $|\hat{\lambda}_2| > |\hat{\lambda}_1|$, um movimento ao longo do eixo u_1 não altera tanto a resposta estimada ($\hat{Y}_0$) como um movimento ao longo do eixo u_2 . A superfície de resposta é mais alongada na direção do eixo u_1 , que apresenta o menor valor $|\hat{\lambda}_1|$.

Para obter a relação entre as variáveis $x_1, x_2, \dots, x_k$ e $u_1, u_2, \dots, u_k$, é preciso calcular a matriz ortogonal M , cujas colunas são formadas pelos autovetores próprios ou característicos normalizados associados aos autovalores da matriz simétrica A .

Autovalores são conhecidos também como raízes características, valores característicos, valores próprios ou latentes. Os autovalores de uma matriz ${}_k A_k$ são as soluções da equação $|A - \lambda I| = 0$ ou $\det(A - \lambda I) = 0$. O determinante é um polinômio de grau k em λ , assim a matriz ${}_k A_k$ gerará k raízes características ou autovalores.

Associado a cada autovalor λ_i de uma matriz quadrada ${}_k A_k$, existe um autovetor m_{vi} , também conhecido como vetor característico, vetor próprio ou vetor latente, cujos elementos satisfazem o sistema de equações homogêneas: $(A - \lambda_i I)m_{vi} = 0_v$. Se λ_i e λ_j são raízes características distintas de uma matriz simétrica B , seus vetores associados m_{vi} e m_{vj} são ortogonais.

$$M = [m_{v1} \quad m_{v2} \quad \dots \quad m_{vk}]$$

Como a matriz M é ortogonal, logo $M' = M^{-1}$ e $M'M = I$. Então $u_v = M'z_v$, em que: $z_v = x_v - x_{v0}$.

11.1.5. Exemplo de um fatorial 3×3

Exercício 11.1.5.1. Seja considerado um ensaio de adubação de uma espécie florestal, instalado em um delineamento inteiramente ao acaso em fatorial 3×3, com duas repetições e dados simulados na tabela 11.1.5.1, com doses de 0, 60, 120 em kg por parcela de fósforo (P) e potássio (K).

Tabela 11.1.5.1. Volume em m³ de madeira por parcela.

Trat.	Rep1	Rep2	Total
P0K0	45,6	49,6	95,2
P0K1	52,5	62,7	164,2
P0K2	69,2	103,5	76,0
P1K0	84,2	80,0	104,0
P1K1	71,0	33,0	205,4
P1K2	83,8	61,0	115,2
P2K0	34,9	41,1	122,1
P2K1	102,2	103,2	172,7
P2K2	62,9	59,2	144,8
Total	606,3	593,3	1.199,6

	P ₀	P ₁	P ₂	Total
K ₀	95,2	104,0	122,1	321,3
K ₁	164,2	205,4	172,7	542,3
K ₂	76,0	115,2	144,8	336,0
Total	335,4	424,6	439,6	1.199,6

$$\hat{\beta}_v = \frac{1}{J} \begin{bmatrix} G/9 \\ \frac{1}{6}(P_2 - P_0) \\ \frac{1}{6}(K_2 - K_0) \\ \frac{1}{6}(P_0 - 2P_1 + P_2) \\ \frac{1}{6}(K_0 - 2K_1 + K_2) \\ \frac{1}{4}[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)] \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1.199,6/9 \\ \frac{1}{6}(439,6 - 335,4) \\ \frac{1}{6}(336,0 - 321,3) \\ \frac{1}{6}(335,4 - 2 \times 424,6 + 439,6) \\ \frac{1}{6}(321,3 - 2 \times 542,3 + 336,0) \\ \frac{1}{4}[(95,2 + 144,8) - (76,0 + 122,1)] \end{bmatrix} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{12} \end{bmatrix} = \begin{bmatrix} 66,644444 \\ 8,683333 \\ 1,225000 \\ -6,183333 \\ -35,608333 \\ 5,2375 \end{bmatrix}$$

$$\hat{Y}_i = 66,644444 + 8,683333x_1 + 1,225x_2 - 6,183333(x_1^2 - \frac{2}{3}) - 35,608333(x_2^2 - \frac{2}{3}) + 5,2375x_1x_2$$

Na tabela 11.1.5.2 estão apresentadas as estimativas dos parâmetros da regressão obtidas pelo SAS, o qual não considera no modelo a inserção da correção "c". Logo, o valor do intercepto será diferente quando usada referida correção no modelo.

Tabela 11.1.5.2. Estimativas dos parâmetros da regressão obtidos pelo SAS.

Parâmetros	G.L.	Estimativas	Erro Padrão	Valor de “t”	Pr > t
Intercepto	1	94,505556	7,201939	13,12	< 0,0001
P	1	8,683333	3,944664	2,20	0,0480
K	1	1,225000	3,944664	0,31	0,7615
P×P	1	-6,183333	6,832359	-0,91	0,3833
K×P	1	5,237500	4,831207	1,08	0,2996
K×K	1	-35,608333	6,832359	-5,21	0,0002

$$SQ_{Total} = \sum_{i=1}^I \sum_{j=1}^J y_{ij}^2 - \frac{G^2}{9J} = 88.554,38 - \frac{1.199,6^2}{18} = 8.607,7044$$

$$SQ_{TRAT} = \frac{Y_v'Y_v}{J} - \frac{G^2}{9J} = \frac{173.777,60}{2} - \frac{1.199,6^2}{18} = 6.942,1244$$

$$SQ_{DREG} = \frac{Y_v'Y_v}{J} - \frac{G^2}{9J} - \left\{ \frac{(P_2 - P_0)^2}{6J} + \frac{(K_2 - K_0)^2}{6J} + \frac{(P_0 - 2P_1 + P_2)^2}{18J} + \frac{(K_0 - 2K_1 + K_2)^2}{18J} + \frac{[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)]}{4J} \right\}$$

$$SQ_{\hat{\alpha}_1} = SQ_{P'} = \frac{(P_2 - P_0)^2}{6J} = \frac{(439,6 - 335,4)^2}{12} = 904,8033$$

$$SQ_{\hat{\alpha}_2} = SQ_{K'} = \frac{(K_2 - K_0)^2}{6J} = \frac{(336,0 - 321,3)^2}{12} = 18,0075$$

$$SQ_{\hat{\alpha}_{11}} = SQ_{P''} = \frac{(P_0 - 2P_1 + P_2)^2}{18J} = \frac{(335,4 - 2 \times 424,6 + 439,6)^2}{36} = 152,9344$$

$$SQ_{\hat{\alpha}_{22}} = SQ_{K''} = \frac{(K_0 - 2K_1 + K_2)^2}{18J} = \frac{(321,3 - 2 \times 542,3 + 336,0)^2}{36} = 5.071,8136$$

$$SQ_{\hat{\alpha}_{12}} = SQ_{(P' \times K')} = \frac{[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)]^2}{4J} = \frac{[(95,2 + 144,8) - (76,0 + 122,1)]^2}{8} = 219,4513$$

$$SQ_{DREG} = 6.942,1244 - (904,8033 + 18,0075 + 152,9344 + 5.071,8136 + 219,4523) = 575,1143$$

$$SQ_{TOTAL} = 88.554,38 - \frac{1.199,6^2}{18} = 8.607,7044$$

Tabela 11.1.5.3. ANOVA da regressão.

F. V.	G.L.	S.Q.	Q.M.	F
Tratamentos	(8)	(6.942,1244)		
P'	1	904,8033	904,8033	4,89 <i>ns</i>
K'	1	18,0075	18,0075	0,10 <i>ns</i>
P''	1	152,9344	152,9344	0,83 <i>ns</i>
K''	1	5.071,8136	5.071,8136	27,41**
$P' \times K'$	1	219,4513	219,4513	1,19 <i>ns</i>
Desvios da regressão	3	575,1143	191,7048	1,04 <i>ns</i>
Resíduo (*)	9	1.665,5800	185,0644	
Total	17	8.607,7044		

Tabela 11.1.5.4. ANOVA da regressão no SAS.

Regressão	G.L.	S.Q.	F	Pr > F
Linear	2	922,810833	2,47	0,1263
Quadrática	2	5224,748056	13,99	0,0007
Interação	1	219,451250	1,18	0,2996
Total	5	6367,010139	6,82	0,0031

$$F_{1;9;0,05} = 5,12 \quad F_{3;9;0,05} = 3,86 \quad F_{1;9;0,01} = 10,56$$

Na tabela 11.1.5.3, o desdobramento está mais detalhado, isto é, considera os efeitos linear e quadrático separadamente para P e K .

Na tabela 11.1.5.4, os resultados obtidos pelo Proc RSREG do SAS são apresentados da seguinte forma:

$$\begin{aligned}
 SQ_{\text{LINEAR}} &= SQ_{P'} + SQ_{K'} = 922,8108 \\
 SQ_{\text{QUADRÁTICO}} &= SQ_{P''} + SQ_{K''} = 5.224,7481 \\
 SQ_{\text{RESÍDUO}} &= SQ_{\text{DREG}} + SQ_{\text{RESÍDUO} (*)} = 575,1143 + 1.665,5800 = 2.240,6943 \\
 SQ_{\text{RESÍDUO}} &= 2.240,6943 \text{ com 12 gl.}
 \end{aligned}$$

Em que:

$$\begin{aligned}
 QM_{\text{RESÍDUO}} &= \frac{2.240,6943}{12} = 186,7245 & F_{\text{LINEAR}} &= \frac{922,810833/2}{186,7245} = 2,47 \\
 F_{\text{QUADRÁTICO}} &= \frac{5.224,748056/2}{186,7245} = 13,99 & F_{\text{INTERAÇÃO}} &= \frac{219,451250}{186,7245} = 1,18
 \end{aligned}$$

Na tabela 11.1.5.2, o valor do intercepto ($\hat{a}_{0(SAS)}$), obtido no SAS, é igual a 94,505556, enquanto, no cálculo do vetor $\hat{\beta}_v$, no exercício 11.1.5.1, o valor foi de 66,644444. No SAS, a análise considerou o modelo 11.1.1.1 sem a inserção da correção $c = \frac{2}{3}$. Por outro lado, no cálculo de $\hat{\beta}_v$, a análise foi realizada de acordo com o modelo 11.1.1.2. A não inserção da correção $c = \frac{2}{3}$ nos termos quadráticos do modelo motivou o valor diferente para a estimativa do parâmetro $\hat{a}_{0(SAS)}$ demonstrado a seguir.

$$\hat{Y}_i = 66,644444 + 8,683333x_1 + 1,225x_2 - 6,183333(x_1^2 - \frac{2}{3}) - 35,608333(x_2^2 - \frac{2}{3}) + 5,2375x_1x_2$$

$$\hat{a}_{0(SAS)} = 66,644444 + (-6,183333 \times -c - 35,608333 \times -c)$$

$$\hat{a}_{0(SAS)} = 66,644444 + (-6,183333 \times -\frac{2}{3} - 35,608333 \times -\frac{2}{3}) = 94,505556$$

Logo:

$$\hat{Y}_i = 94,505556 + 8,683333x_1 + 1,225x_2 - 6,183333x_1^2 - 35,608333x_2^2 + 5,2375x_1x_2$$

a) Cálculo das doses economicamente aconselháveis:

$$\hat{Y}_i = 94,505556 + 8,683333x_1 + 1,225x_2 - 6,183333x_1^2 - 35,608333x_2^2 + 5,2375x_1x_2$$

$$q_1 = q_2 = 60 \quad \frac{t_1}{w} = \frac{t_2}{w} = 0,10$$

$$\begin{cases} 2\hat{a}_{11}x_1^* + \hat{a}_{12}x_2^* = \frac{t_1q_1}{w} - \hat{a}_1 \\ \hat{a}_{12}x_1^* + 2\hat{a}_{22}x_2^* = \frac{t_2q_2}{w} - \hat{a}_2 \end{cases}$$

$$\begin{cases} 2 \times -6,183333x_1^* + 5,2375x_2^* = 0,10 \times 60 - 8,683333 \\ 5,2375x_1^* + 2 \times -35,608333x_2^* = 0,10 \times 60 - 1,225 \end{cases}$$

$$\begin{cases} -12,366667x_1^* + 5,2375x_2^* = -2,683333 \\ 5,2375x_1^* - 71,216666x_2^* = 4,775 \end{cases}$$

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} -12,366667 & 5,2375 \\ 5,2375 & -71,216666 \end{bmatrix}^{-1} \times \begin{bmatrix} -2,683333 \\ 4,775 \end{bmatrix} = \begin{bmatrix} 0,194647 \\ -0,052734 \end{bmatrix}$$

$$0,194647 = \frac{X_1^* - 60}{60} \quad X_1^* = 71,67882$$

$$-0,052734 = \frac{X_2^* - 60}{60} \quad X_2^* = 56,83596$$

Por outro lado, a produção no ponto economicamente recomendável, por parcela, corresponde a $\hat{Y}_0^* = \hat{Y}(X_1^*, X_2^*) = 95,744088 \text{ m}^3$. O ponto $[X_1^*, X_2^*] = [71,67882 \ 56,83596]$ encontra-se dentro do espaço experimental pesquisado, isto é, (0, 60, 120 kg).

b) Cálculo do ponto estacionário sem considerar o aspecto econômico:

$$\begin{cases} -12,366667x_{01} + 5,2375x_{02} = -8,683333 \\ 5,2375x_{01} - 71,216666x_{02} = -1,225 \end{cases}$$

$$\begin{bmatrix} x_{01} \\ x_{02} \end{bmatrix} = \begin{bmatrix} -12,366667 & 5,2375 \\ 5,2375 & -71,216666 \end{bmatrix}^{-1} \times \begin{bmatrix} -8,683333 \\ -1,225 \end{bmatrix} = \begin{bmatrix} 0,732248 \\ 0,071053 \end{bmatrix}$$

$$0,732248 = \frac{X_{01} - 60}{60} \quad X_{01} = 103,93488$$

$$0,071053 = \frac{X_{02} - 60}{60} \quad X_{02} = 64,263174$$

A produção no ponto estacionário, sem considerar o aspecto econômico, por parcela, corresponde a $\hat{Y}_0 = \hat{Y}(X_{01}, X_{02}) = 97,728254 \text{ m}^3$. O ponto de máximo estacionário (X_{01}, X_{02}) , sem considerar a análise econômica, está dentro do espaço amostral experimental, ou seja, (0, 60, 120 kg).

c) Verificação se (X_{01}, X_{02}) é um ponto de máximo:

$$d^2 RL = \begin{bmatrix} dx_1 & dx_2 \end{bmatrix} \begin{bmatrix} \hat{a}_{11} & \frac{\hat{a}_{12}}{2} \\ \frac{\hat{a}_{12}}{2} & \hat{a}_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix}$$

Seja considerada a matriz Hessiana A :

$$A = \begin{bmatrix} \hat{a}_{11} & \frac{\hat{a}_{12}}{2} \\ \frac{\hat{a}_{12}}{2} & \hat{a}_{22} \end{bmatrix} = \begin{bmatrix} -6,183333 & 2,61875 \\ 2,61875 & -35,608333 \end{bmatrix}$$

Quando os termos da diagonal principal da matriz Hessiana (A) são negativos, há indicação da ocorrência de um ponto de máximo. No caso, se ambos forem positivos, então poderia ser um ponto de mínimo, entretanto precisa ser confirmado. Por meio da matriz Hessiana, existem dois processos para estudar e confirmar a natureza do ponto obtido:

a) Obter os autovalores ou raízes próprias ou características da matriz Hessiana. Quando todos são positivos, a natureza do ponto é de mínimo. No caso de serem todos negativos, o ponto é de máximo. Quando existem autovalores positivos e negativos, o ponto é de sela. Esse processo é utilizado pelo programa “RSREG” do SAS.

b) Obter, por meio de operações sobre as linhas da matriz Hessiana, uma matriz triangular, ou seja, só com termos nulos abaixo da diagonal principal. No caso de ocorrerem somente termos positivos na diagonal, a natureza do ponto é de mínimo. Entretanto, se todos forem negativos, o ponto será de máximo. Por outro lado, se ocorrerem termos positivos e negativos, o ponto será de sela.

De:

$$\det(A - \lambda I) = 0$$

$$\det \left\{ \begin{bmatrix} -6,183333 & 2,61875 \\ 2,61875 & -35,608333 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} = 0$$

$$\det \begin{bmatrix} -6,183333 - \lambda & 2,61875 \\ 2,61875 & -35,608333 - \lambda \end{bmatrix} = 0$$

$$(-6,183333 - \lambda)(-35,608333 - \lambda) - 2,61875^2 = 0$$

$$\lambda^2 + 41,791666\lambda + 213,320329 = 0$$

$$\lambda_i = \frac{-41,791666 \pm \sqrt{41,791666^2 - 4 \times 213,320329}}{2}$$

$$\hat{\lambda}_1 = \frac{-41,791666 + 29,887490}{2} = -5,952088$$

$$\hat{\lambda}_2 = \frac{-41,791666 - 29,887490}{2} = -35,839578$$

Resultando: $\hat{\lambda}_1 = -5,952089$ e $\hat{\lambda}_2 = -35,83958$

Como os autovalores $\hat{\lambda}_1$ e $\hat{\lambda}_2$ são negativos, a matriz Hessiana A é definida negativa. Logo (X_{01}, X_{02}) é um ponto de máximo, tal que $X_{01} = 103,93488$ e $X_{02} = 64,263174$.

Matriz M com os autovetores correspondentes aos autovalores $\hat{\lambda}_1$, $\hat{\lambda}_2$ associados à matriz A .

De $\det(A - \lambda I) = 0$, surgem os autovalores: $\hat{\lambda}_1 = -5,952089$ e $\hat{\lambda}_2 = -35,83958$. A matriz dos autovetores $M = [a_{v1} \ a_{v2}]$ é obtida pelo sistema de equações $(A - \lambda_i I)a_v = 0_v$.

O programa **Prog34**, em anexo no apêndice B, desenvolvido no SAS (Proc IML), efetua os cálculos para obtenção dos autovalores e autovetores.

$$M = [a_{v1} \ a_{v2}] = \begin{bmatrix} 0,9961239 & -0,0879612 \\ 0,0879612 & 0,9961239 \end{bmatrix}$$

Da definição da forma canônica $\hat{Y} = \hat{Y}_0 + \hat{\lambda}_1 u_1^2 + \hat{\lambda}_2 u_2^2$, tal que o valor 97,728254 é a produção no ponto estacionário:

$$\hat{Y} = 97,728254 - 5,952089 u_1^2 - 35,83958 u_2^2$$

Como $|\hat{\lambda}_2| > |\hat{\lambda}_1|$, então ocorre no eixo (u_2) que corresponde ao potássio, um decréscimo do ponto de máximo em direção às extremidades. Esse decréscimo é maior do que o ocorrente no eixo (u_1) do fósforo.

A matriz M é ortogonal, logo $M^{-1} = M'$, então $MM' = I$. Seja $u_v = M'z_v$, em que $z_v = x_v - x_{v0}$.

$$M = \begin{bmatrix} 0,9961239 & 0,0879612 \\ -0,0879612 & 0,9961239 \end{bmatrix}$$

$$x_{v0} = \begin{bmatrix} 0,732248 \\ 0,071053 \end{bmatrix}$$

$$z_v = x_v - x_{v0} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0,732248 \\ 0,071053 \end{bmatrix}$$

Considerando a relação entre as variáveis $x'_v = [x_1, x_2]$ e $u'_v = [u_1, u_2]$:

$$u_v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = M'z_v = \begin{bmatrix} 0,9961239 & 0,0879612 \\ -0,0879612 & 0,9961239 \end{bmatrix} \times \begin{bmatrix} x_1 - 0,732248 \\ x_2 - 0,071053 \end{bmatrix}$$

Algumas comprovações importantes:

Como $\hat{\lambda}_1 = -5,952089$ e $\hat{\lambda}_2 = -35,83958$, então:

$$D = \begin{bmatrix} -5,952089 & 0 \\ 0 & -35,83958 \end{bmatrix}$$

$$A = [MDM'] = \begin{bmatrix} 0,9961239 & -0,0879612 \\ 0,0879612 & 0,9961239 \end{bmatrix} \times \begin{bmatrix} -5,952089 & 0 \\ 0 & -35,83958 \end{bmatrix} \times \begin{bmatrix} 0,9961239 & 0,0879612 \\ -0,0879612 & 0,9961239 \end{bmatrix}$$

$$A = \begin{bmatrix} -6,183333 & 2,61875 \\ 2,61875 & -35,608333 \end{bmatrix}$$

Assim como:

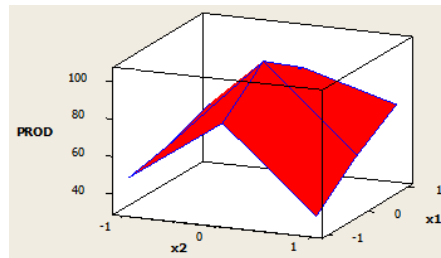
$$D = [M'AM] = \begin{bmatrix} 0,9961239 & 0,0879612 \\ -0,0879612 & 0,9961239 \end{bmatrix} \times \begin{bmatrix} -6,183333 & 2,61875 \\ 2,61875 & -35,608333 \end{bmatrix} \times \begin{bmatrix} 0,9961239 & -0,0879612 \\ 0,0879612 & 0,9961239 \end{bmatrix}$$

$$D = \begin{bmatrix} -5,952089 & 0 \\ 0 & -35,83958 \end{bmatrix}$$

Tabela 11.1.5.5. Valores críticos e suas naturezas no SAS.

Fatores	Valores Críticos Codificados	Autovetores		
		Autovalores	P	K
P	0,732248	-5,952089	0,996124	0,087961
K	0,071053	-35,839578	-0,087961	0,996124
Valor estimado no ponto estacionário: 97,728254		O ponto estacionário é um máximo		

Pode-se perceber, na figura 11.1.5.1, a localização do ponto estacionário de máximo $(X_{01}, X_{02}) = (P, K) = (0,732248, 0,071053)$, que corresponde ao valor da produção máxima ($\hat{Y}_0 = 97,728254 \text{ m}^3$ por parcela).

**Figura 11.1.5.1.** Superfície de resposta fatorial 3×3 do exercício 11.1.5.1.

O programa **Prog33**, em anexo no apêndice B, desenvolvido no SAS, executa a análise de regressão (tabelas 11.1.5.2 e 11.1.5.4), assim como apresenta os valores do ponto crítico (X_{01}, X_{02}) , demonstrando ser um ponto estacionário de máximo (tabela 11.1.5.5).

11.2. ANÁLISE DE SUPERFÍCIE DE RESPOSTA DE UM ENSAIO FATORIAL $3 \times 3 \times 3$

11.2.1. Estimativas dos parâmetros da regressão

Dado um experimento de adubação em um delineamento inteiramente ao acaso em fatorial $3 \times 3 \times 3$.

Seja considerado o modelo quadrático de segunda ordem:

$$Y_i = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + a_{11}X_1^2 + a_{22}X_2^2 + a_{33}X_3^2 + a_{12}X_1X_2 + a_{13}X_1X_3 + a_{23}X_2X_3 + e_i \quad 11.2.1.1.$$

Dada a transformação:

$$x_1 = \frac{X_1 - \bar{X}_1}{q_1} \quad x_2 = \frac{X_2 - \bar{X}_2}{q_2} \quad x_3 = \frac{X_3 - \bar{X}_3}{q_3}$$

$$Y_i = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + e_i \quad 11.2.1.2$$

Utilizando a correção “c” nos termos quadráticos:

$$Y_i = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{11}(x_1^2 - c) + a_{22}(x_2^2 - c) + a_{33}(x_3^2 - c) + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + e_i \quad 11.2.1.3$$

Usando o mesmo procedimento para a obtenção da constante (c) que foi adotado no caso do fatorial 3×3, resulta para o fatorial 3×3×3 (NPK):

$$18(1-c) - 9c = 18 - 18c - 9c = 0 \quad 27c = 18 \quad c = \frac{2}{3}$$

Resultando em:

$$\hat{\beta}_v = (X'X)^{-1} X'Y_v = \frac{1}{J} \begin{bmatrix} G/27 \\ \frac{1}{18}(N_2 - N_0) \\ \frac{1}{18}(P_2 - P_0) \\ \frac{1}{18}(K_2 - K_0) \\ \frac{1}{18}(N_0 - 2N_1 + N_2) \\ \frac{1}{18}(P_0 - 2P_1 + P_2) \\ \frac{1}{18}(K_0 - 2K_1 + K_2) \\ \frac{1}{12}[(N_0P_0 + N_2P_2) - (N_0P_2 + N_2P_0)] \\ \frac{1}{12}[(N_0K_0 + N_2K_2) - (N_0K_2 + N_2K_0)] \\ \frac{1}{12}[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)] \end{bmatrix} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{33} \\ \hat{a}_{12} \\ \hat{a}_{13} \\ \hat{a}_{23} \end{bmatrix}$$

Em que a soma de quadrados dos desvios da regressão (SQ_{DREG}) é dada pela fórmula:

$$SQ_{DREG} = \frac{Y_v'Y_v}{J} - \frac{G^2}{27J} - \left\{ \frac{(N_2 - N_0)^2}{18J} + \frac{(P_2 - P_0)^2}{18J} + \frac{(K_2 - K_0)^2}{18J} + \frac{(N_0 - 2N_1 + N_2)^2}{54J} + \frac{(P_0 - 2P_1 + P_2)^2}{54J} + \frac{(K_0 - 2K_1 + K_2)^2}{54J} \right.$$

$$\left. + \frac{[(N_0P_0 + N_2P_2) - (N_0P_2 + N_2P_0)]^2}{12J} + \frac{[(N_0K_0 + N_2K_2) - (N_0K_2 + N_2K_0)]^2}{12J} + \frac{[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)]^2}{12J} \right\}$$

$$SQ_{DREG} = SQ_{TRAT} - [SQ_{N'} + SQ_{P'} + SQ_{K'} + SQ_{N''} + SQ_{P''} + SQ_{K''} + SQ(N' \times P') + SQ(N' \times K') + SQ(P' \times K')]$$

O vetor é formado pelos valores dos totais dos vinte e sete tratamentos.

11.2.2. Doses econômicas aconselháveis

$$\hat{Y}_i = \hat{a}_0 + \hat{a}_1 x_1 + \hat{a}_2 x_2 + \hat{a}_3 x_3 + \hat{a}_{11} x_1^2 + \hat{a}_{22} x_2^2 + \hat{a}_{33} x_3^2 + \hat{a}_{12} x_1 x_2 + \hat{a}_{13} x_1 x_3 + \hat{a}_{23} x_2 x_3$$

$$\begin{aligned} RL &= w\hat{Y}_i - cf - t_1 q_1 x_1 - t_2 q_2 x_2 - t_3 q_3 x_3 \\ \frac{RL}{w} &= \hat{Y}_i - \frac{cf}{w} - \frac{t_1 q_1 x_1}{w} - \frac{t_2 q_2 x_2}{w} - \frac{t_3 q_3 x_3}{w} \\ \frac{dRL}{w} &= d\hat{Y}_i - \frac{t_1 q_1}{w} dx_1 - \frac{t_2 q_2}{w} dx_2 - \frac{t_3 q_3}{w} dx_3 \end{aligned}$$

$$\begin{aligned} d\hat{Y}_i &= (\hat{a}_1 + 2\hat{a}_{11}x_1 + \hat{a}_{12}x_2 + \hat{a}_{13}x_3)dx_1 + (\hat{a}_2 + 2\hat{a}_{22}x_2 + \hat{a}_{12}x_1 + \hat{a}_{23}x_3)dx_2 + \\ &(\hat{a}_3 + 2\hat{a}_{33}x_3 + \hat{a}_{13}x_1 + \hat{a}_{23}x_2)dx_3 \end{aligned}$$

$$\begin{aligned} \frac{dL}{w} &= (\hat{a}_1 + 2\hat{a}_{11}x_1 + \hat{a}_{12}x_2 + \hat{a}_{13}x_3 - \frac{t_1 q_1}{w})dx_1 + (\hat{a}_2 + 2\hat{a}_{22}x_2 + \hat{a}_{12}x_1 + \hat{a}_{23}x_3 - \frac{t_2 q_2}{w})dx_2 + \\ &(\hat{a}_3 + 2\hat{a}_{33}x_3 + \hat{a}_{13}x_1 + \hat{a}_{23}x_2 - \frac{t_3 q_3}{w})dx_3 \end{aligned}$$

$$\begin{aligned} \hat{a}_1 + 2\hat{a}_{11}x_1^* + \hat{a}_{12}x_2^* + \hat{a}_{13}x_3^* &= \frac{t_1 q_1}{w} \\ \hat{a}_2 + 2\hat{a}_{22}x_2^* + \hat{a}_{12}x_1^* + \hat{a}_{23}x_3^* &= \frac{t_2 q_2}{w} \\ \hat{a}_3 + 2\hat{a}_{33}x_3^* + \hat{a}_{13}x_1^* + \hat{a}_{23}x_2^* &= \frac{t_3 q_3}{w} \end{aligned}$$

Concluindo: (x_1^*, x_2^*, x_3^*) é a dose economicamente aconselhável.

$$\begin{cases} 2\hat{a}_{11}x_1^* + \hat{a}_{12}x_2^* + \hat{a}_{13}x_3^* = \frac{t_1 q_1}{w} - \hat{a}_1 \\ \hat{a}_{12}x_1^* + 2\hat{a}_{22}x_2^* + \hat{a}_{23}x_3^* = \frac{t_2 q_2}{w} - \hat{a}_2 \\ \hat{a}_{13}x_1^* + \hat{a}_{23}x_2^* + 2\hat{a}_{33}x_3^* = \frac{t_3 q_3}{w} - \hat{a}_3 \end{cases}$$

Obtenção da diferencial de segunda ordem para estudar a natureza do ponto crítico:

$$\begin{aligned} \frac{dL}{w} &= (\hat{a}_1 + 2\hat{a}_{11}x_1 + \hat{a}_{12}x_2 + \hat{a}_{13}x_3 - \frac{t_1 q_1}{w})dx_1 + (\hat{a}_2 + 2\hat{a}_{22}x_2 + \hat{a}_{12}x_1 + \hat{a}_{23}x_3 \\ &- \frac{t_2 q_2}{w})dx_2 + (\hat{a}_3 + 2\hat{a}_{33}x_3 + \hat{a}_{13}x_1 + \hat{a}_{23}x_2 - \frac{t_3 q_3}{w})dx_3 \\ \frac{d^2 L}{w} &= 2\hat{a}_{11}d^2 x_1 + 2\hat{a}_{22}d^2 x_2 + 2\hat{a}_{33}d^2 x_3 + \hat{a}_{12}dx_1 dx_2 + \hat{a}_{12}dx_1 dx_2 + \\ &\hat{a}_{13}dx_1 dx_3 + \hat{a}_{13}dx_1 dx_3 + \hat{a}_{23}dx_2 dx_3 + \hat{a}_{23}dx_2 dx_3 \end{aligned}$$

$$\frac{d^2L}{dw} = 2\hat{a}_{11}d^2x_1 + 2\hat{a}_{22}d^2x_2 + 2\hat{a}_{33}d^2x_3 + 2\hat{a}_{12}dx_1dx_2 + 2\hat{a}_{13}dx_1dx_3 + 2\hat{a}_{23}dx_2dx_3$$

$$\frac{d^2L}{2w} = \begin{bmatrix} dx_1 & dx_2 & dx_3 \end{bmatrix} \begin{bmatrix} \hat{a}_{11} & \frac{\hat{a}_{12}}{2} & \frac{\hat{a}_{13}}{2} \\ \frac{\hat{a}_{12}}{2} & \hat{a}_{22} & \frac{\hat{a}_{23}}{2} \\ \frac{\hat{a}_{13}}{2} & \frac{\hat{a}_{23}}{2} & \hat{a}_{33} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \\ dx_3 \end{bmatrix}$$

Em que a matriz A é denominada de matriz Hessiana.

$$A = \begin{bmatrix} \hat{a}_{11} & \frac{\hat{a}_{12}}{2} & \frac{\hat{a}_{13}}{2} \\ \frac{\hat{a}_{12}}{2} & \hat{a}_{22} & \frac{\hat{a}_{23}}{2} \\ \frac{\hat{a}_{13}}{2} & \frac{\hat{a}_{23}}{2} & \hat{a}_{33} \end{bmatrix}$$

- a) Se a matriz Hessiana for definida negativa, $x^* = (x_1^*, x_2^*, x_3^*)$ será um ponto de máximo;
 b) Se a matriz Hessiana for definida positiva, $x^* = (x_1^*, x_2^*, x_3^*)$ será um ponto de mínimo;
 c) Se a matriz Hessiana A for não definida, $x^* = (x_1^*, x_2^*, x_3^*)$ será um ponto de sela (*Saddle point*), ou seja, nem máximo e nem mínimo.

11.2.3. Um exemplo de aplicação de um fatorial 3×3×3

Exercício 11.2.3.1. Seja um ensaio de adubação em uma espécie florestal, com duas repetições ($J = 2$) e as doses de 0-60-120 em kg por parcela de Nitrogênio (N), Fósforo (P) e Potássio (K), instalado em um ensaio inteiramente ao acaso em um fatorial 3×3×3.

Tabela 11.2.3.1. Volume em m³ por parcela de madeira (dados simulados).

Trat.	Rep.	Prod.	Trat.	Rep.	Prod.	Trat.	Rep.	Prod.
N0P0K0	1	45,6	N0P0K1	1	84,2	N0P0K2	1	34,9
N0P1K2	1	52,5	N0P1K0	1	71,0	N0P1K1	1	92,2
N0P2K1	1	69,2	N0P2P2	1	83,8	N0P2K0	1	62,9
N1P0K1	1	57,0	N1P0K2	1	43,3	N1P0K0	1	35,9
N1P1K0	1	66,3	N1P1K1	1	98,2	N1P1K2	1	88,2
N1P2K2	1	71,3	N1P2K0	1	69,2	N1P2K1	1	107,7
N2P0K2	1	44,6	N2P0K0	1	69,8	N2P0K1	1	60,9
N2P1K1	1	43,3	N2P1K2	1	49,3	N2P1K0	1	40,7
N2P2K0	1	60,4	N2P2K1	1	83,2	N2P2K2	1	55,5
N0P0K0	2	36,1	N0P0K1	2	58,2	N0P0K2	2	29,9
N0P1K2	2	45,6	N0P1K0	2	24,0	N0P1K1	2	61,5
N0P2K1	2	75,3	N0P2P2	2	44,4	N0P2K0	2	43,1
N1P0K1	2	42,2	N1P0K2	2	29,7	N1P0K0	2	32,2
N1P1K0	2	85,6	N1P1K1	2	97,7	N1P1K2	2	72,6
N1P2K2	2	75,1	N1P2K0	2	75,8	N1P2K1	2	90,0
N2P0K2	2	43,4	N2P0K0	2	37,6	N2P0K1	2	39,6
N2P1K1	2	65,5	N2P1K2	2	46,8	N2P1K0	2	82,9
N2P2K0	2	56,7	N2P2K1	2	50,7	N2P2K2	2	60,4

Quadro, considerando as 27 combinações:

	P0	P1	P2	Total		K0	K1	K2	Total
N0	288,9	346,8	378,7	1.014,4	N0	282,7	440,6	291,1	1.014,4
N1	240,3	508,6	489,1	1.238,0	N1	365,0	492,8	380,2	1.238,0
N2	295,9	328,5	366,9	991,3	N2	348,1	343,2	300,0	991,3
Total	825,1	1.183,9	1.234,7	3.243,7	Total	995,8	1.276,6	971,3	3.243,7

	K0	K1	K2	Total
P0	257,2	342,1	225,8	825,1
P1	370,5	458,4	355,0	1.183,9
P2	368,1	476,1	390,5	1.234,7
Total	995,8	1.276,6	971,3	3.243,7

a) Estimativas dos parâmetros

$$\hat{\beta} = (X'X)^{-1}X'Y_v = \frac{1}{J} \begin{bmatrix} G/27 \\ \frac{1}{18}(N_2 - N_0) \\ \frac{1}{18}(P_2 - P_0) \\ \frac{1}{18}(K_2 - K_0) \\ \frac{1}{18}(N_0 - 2N_1 + N_2) \\ \frac{1}{18}(P_0 - 2P_1 + P_2) \\ \frac{1}{18}(K_0 - 2K_1 + K_2) \\ \frac{1}{12}[(N_0P_0 + N_2P_2) - (N_0P_2 + N_2P_0)] \\ \frac{1}{12}[(N_0K_0 + N_2K_2) - (N_0K_2 + N_2K_0)] \\ \frac{1}{12}[(P_0K_0 + P_2K_2) - (P_0K_2 + P_2K_0)] \end{bmatrix} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{33} \\ \hat{a}_{12} \\ \hat{a}_{13} \\ \hat{a}_{23} \end{bmatrix} = \begin{bmatrix} 60,068519 \\ -0,641667 \\ 11,377778 \\ -0,680556 \\ -13,063889 \\ -8,555556 \\ -16,280556 \\ -0,783333 \\ -2,354167 \\ 2,241667 \end{bmatrix}$$

Tabela 11.2.3.2. Estimativas dos parâmetros pelo SAS.

Parâmetros	G.L.	Estimativas	Erro Padrão	Valor “t”	Pr > t
Intercepto	1	85,335185	5,674098	15,04	< 0,0001
N	1	-0,641667	2,626597	-0,24	0,8081
P	1	11,377778	2,626597	4,33	< 0,0001
K	1	-0,680556	2,626597	-0,26	0,7968
N×N	1	-13,063889	4,549399	-2,87	0,0063
P×N	1	-0,783333	3,216911	-0,24	0,8087
P×P	1	-8,555556	4,549399	-1,88	0,0667
K×N	1	-2,354167	3,216911	-0,73	0,4682
K×P	1	2,241667	3,216911	0,70	0,4896
K×K	1	-16,280556	4,549399	-3,58	0,0009

A tabela 11.2.3.2 apresenta as estimativas dos parâmetros da regressão pelo SAS, conforme o **Prog35**, em anexo no apêndice B. O programa não usa a correção “c”.

$$\hat{Y}_i = 60,068519 - 0,641667x_1 + 11,377778x_2 - 0,680556x_3 - 13,063889(x_1^2 - \frac{2}{3}) - 8,555556(x_2^2 - \frac{2}{3}) - 16,280556(x_3^2 - \frac{2}{3}) - 0,783333x_1x_2 - 2,354167x_1x_3 + 2,241667x_2x_3$$

Tal que:

$$\hat{Y}_i = 85,335185 - 0,641667x_1 + 11,377778x_2 - 0,680556x_3 - 13,063889x_1^2 - 8,555556x_2^2 - 16,280556x_3^2 - 0,783333x_1x_2 - 2,354167x_1x_3 + 2,241667x_2x_3$$

b) Cálculo das doses economicamente aconselháveis:

$$\begin{cases} 2\hat{a}_{11}x_1^* + \hat{a}_{12}x_2^* + \hat{a}_{13}x_3^* = \frac{t_1q_1}{w} - \hat{a}_1 \\ \hat{a}_{12}x_1^* + 2\hat{a}_{22}x_2^* + \hat{a}_{23}x_3^* = \frac{t_2q_2}{w} - \hat{a}_2 \\ \hat{a}_{13}x_1^* + \hat{a}_{23}x_2^* + 2\hat{a}_{33}x_3^* = \frac{t_3q_3}{w} - \hat{a}_3 \end{cases}$$

Considerando: $q_1 = q_2 = q_3 = 60$ e $\frac{t_1}{w} = \frac{t_2}{w} = \frac{t_3}{w} = 0,10$:

$$\begin{cases} 2 \times -13,063889x_1^* - 0,783333x_2^* - 2,354167x_3^* = 0,10 \times 60 - (-0,641667) \\ -0,783333x_1^* + 2 \times -8,555556x_2^* + 2,241667x_3^* = 0,10 \times 60 - 11,377778 \\ -2,354167x_1^* + 2,241667x_2^* + 2 \times -16,280556x_3^* = 0,10 \times 60 - (-0,680556) \end{cases}$$

$$\begin{cases} -26,127778x_1^* - 0,783333x_2^* - 2,354167x_3^* = 6,641667 \\ -0,783333x_1^* - 17,111112x_2^* + 2,241667x_3^* = -5,377778 \\ -2,354167x_1^* + 2,241667x_2^* - 32,561112x_3^* = 6,680556 \end{cases}$$

$$\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} -26,127778 & -0,783333 & -2,354167 \\ -0,783333 & -17,111112 & 2,241667 \\ -2,354167 & 2,241667 & -32,561112 \end{bmatrix}^{-1} \times \begin{bmatrix} 6,641667 \\ -5,377778 \\ 6,680556 \end{bmatrix}$$

$$\begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} -0,248326 \\ 0,303868 \\ -0,166296 \end{bmatrix}$$

$$-0,248326 = \frac{X_1^* - 60}{60}$$

$$X_1^* = 45,10044$$

$$0,303868 = \frac{X_2^* - 60}{60} \quad X_2^* = 78,23208 \quad -0,166296 = \frac{X_3^* - 60}{60} \quad X_3^* = 50,02224$$

O ponto economicamente aconselhável deverá apresentar a seguinte produção de:

$$Y^* = 86,867853 \text{ m}^3 \text{ por parcela}$$

O ponto $X^* = [45,10044 \ 78,23208 \ 50,02224]$ corresponde a uma produção máxima estimada de $\hat{Y}_0 = 86,867853 \text{ m}^3$, assim como pertence ao espaço experimental pesquisado.

c) Cálculo das doses não considerando a análise econômica aconselhável:

$$\hat{Y}_i = 85,335185 - 0,641667x_1 + 11,377778x_2 - 0,680556x_3 - 13,063889x_1^2 - 8,555556x_2^2 - 16,280556x_3^2 - 0,783333x_1x_2 - 2,354167x_1x_3 + 2,241667x_2x_3$$

$$\begin{cases} 2\hat{a}_{11}x_{01} + \hat{a}_{12}x_{02} + \hat{a}_{13}x_{03} = -\hat{a}_1 \\ \hat{a}_{12}x_{01} + 2\hat{a}_{22}x_{02} + \hat{a}_{23}x_{03} = -\hat{a}_2 \\ \hat{a}_{13}x_{01} + \hat{a}_{23}x_{02} + 2\hat{a}_{33}x_{03} = -\hat{a}_3 \end{cases}$$

$$\begin{cases} 2 \times -13,063889x_{01} - 0,783333x_{02} - 2,354167x_{03} = -(-0,641667) \\ -0,783333x_{01} + 2 \times -8,555556x_{02} + 2,241667x_{03} = -11,377778 \\ -2,354167x_{01} + 2,241667x_{02} + 2 \times -16,280556x_{03} = -(-0,680556) \end{cases}$$

$$\begin{cases} -26,127778x_{01} - 0,783333x_{02} - 2,354167x_{03} = 0,641667 \\ -0,783333x_{01} - 17,111112x_{02} + 2,241667x_{03} = -11,377778 \\ -2,354167x_{01} + 2,241667x_{02} - 32,561112x_{03} = 0,680556 \end{cases}$$

$$\begin{bmatrix} x_{01} \\ x_{02} \\ x_{03} \end{bmatrix} = \begin{bmatrix} -26,127778 & -0,783333 & -2,354167 \\ -0,783333 & -17,111112 & 2,241667 \\ -2,354167 & 2,241667 & -32,561112 \end{bmatrix}^{-1} \times \begin{bmatrix} 0,641667 \\ -11,377778 \\ 0,680556 \end{bmatrix}$$

$$\begin{bmatrix} x_{01} \\ x_{02} \\ x_{03} \end{bmatrix} = \begin{bmatrix} -0,047258 \\ 0,670859 \\ 0,028701 \end{bmatrix}$$

$$-0,047258 = \frac{X_{01} - 60}{60} \quad X_{01} = 57,16452$$

$$0,670859 = \frac{X_{02} - 60}{60} \quad X_{02} = 100,25154 \quad 0,028701 = \frac{X_{03} - 60}{60} \quad X_{03} = 61,72206$$

O ponto $X'_{v0} = [57,16452 \ 100,25154 \ 61,72206]$ corresponde a uma produção máxima estimada de $\hat{Y}_0 = 89,157020 \text{ m}^3$ por parcela, assim como pertence ao espaço experimental pesquisado (0, 60, 120 kg).

d) Verificação da natureza do ponto crítico:

$$\det(A - I\lambda) = 0$$

$$A = \begin{bmatrix} \hat{a}_{11} & \frac{\hat{a}_{12}}{2} & \frac{\hat{a}_{13}}{2} \\ \frac{\hat{a}_{12}}{2} & \hat{a}_{22} & \frac{\hat{a}_{23}}{2} \\ \frac{\hat{a}_{13}}{2} & \frac{\hat{a}_{23}}{2} & \hat{a}_{33} \end{bmatrix} = \begin{bmatrix} -13,063889 & -0,3916665 & -1,1770835 \\ -0,3916665 & -8,555556 & 1,1208335 \\ -1,1770835 & 1,1208335 & -16,280556 \end{bmatrix}$$

$$\det \begin{bmatrix} -13,063889 & -0,3916665 & -1,1770835 \\ -0,3916665 & -8,555556 & 1,1208335 \\ -1,1770835 & 1,1208335 & -16,280556 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 0$$

$$\hat{\lambda}_1 = -8,32938 \quad \hat{\lambda}_2 = -12,795305 \quad \hat{\lambda}_3 = -16,775316$$

Considerando que a matriz Hessiana A possui todos os autovalores negativos, conclui-se que A é definida negativa. Logo $X'_{v0} = [57,16452 \quad 100,25154 \quad 61,72206]$ é um ponto de máximo. O ponto de máximo está dentro do espaço amostral (0, 60, 120) kg por parcela, apresentando a produção máxima de $\hat{Y}_0 = 89,157020 \text{ m}^3$ por parcela.

O programa SAS mostra os valores do ponto crítico (X_{01}, X_{02}, X_{03}), sem considerar a análise econômica, bem como avalia a sua natureza. A tabela 11.2.3.3 apresenta os resultados e a conclusão que o ponto crítico obtido é um ponto estacionário de máximo.

Tabela 11.2.3.3. Valores críticos e suas naturezas no SAS.

Fatores	Valores Críticos Codificados	Autovalores	Autovetores		
			N	P	K
N	-0,047258	-8,329380	-0,11988	0,980462	0,155958
P	0,670859	-12,795305	0,949746	0,159015	-0,269622
K	0,028701	-16,775316	0,289154	-0,115797	0,950253
Valor estimado do ponto estacionário: 89,157020		O ponto estacionário é um máximo			

Da definição da forma canônica $\hat{Y} = \hat{Y}_0 + \hat{\lambda}_1 u_1^2 + \hat{\lambda}_2 u_2^2 + \hat{\lambda}_3 u_3^2$, resulta:

$$\hat{\lambda}_1 = -8,32938 \quad \hat{\lambda}_2 = -12,795305 \quad \hat{\lambda}_3 = -16,775316$$

Então:

$$\hat{Y} = 89,157020 - 8,32938 u_1^2 - 12,795305 u_2^2 - 16,775316 u_3^2$$

Seja a matriz M dos autovetores normalizados:

$$M = \begin{bmatrix} -0,11988 & 0,949746 & 0,289154 \\ 0,980462 & 0,159015 & -0,115797 \\ 0,155958 & -0,269622 & 0,950253 \end{bmatrix}$$

A matriz M é ortogonal, logo $M^{-1} = M'$, então $MM' = I$. Seja $u_v = M'z_v$, em que $z_v = x_v - x_{v0}$.

$$x_{v0} = \begin{bmatrix} -0,047258 \\ 0,670859 \\ 0,028701 \end{bmatrix} \quad z_v = x_v - x_{v0} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} -0,047258 \\ 0,670859 \\ 0,028701 \end{bmatrix}$$

A relação entre as variáveis $x'_v = [x_1, x_2, x_3]$ e $u'_v = [u_1, u_2, u_3]$:

$$u_v = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = M'z_v = \begin{bmatrix} -0,11988 & 0,980462 & 0,155958 \\ 0,949746 & 0,159015 & -0,269622 \\ 0,289154 & -0,115797 & 0,950253 \end{bmatrix} \times \begin{bmatrix} x_1 + 0,047258 \\ x_2 - 0,670859 \\ x_3 - 0,028701 \end{bmatrix}$$

O fósforo e o potássio, figura 11.2.3.1, apresentaram uma maior contribuição na construção da variável canônica, pois apresentaram os dois maiores autovalores, $\hat{\lambda}_2 = -12,795305$ e $\hat{\lambda}_3 = -16,775316$ ($|\lambda_2|$ e $|\lambda_3|$). Referida figura mostra uma visão da superfície de resposta, tal que a produção máxima é $\hat{Y}_0 = 89,157020 \text{ m}^3$ por parcela.

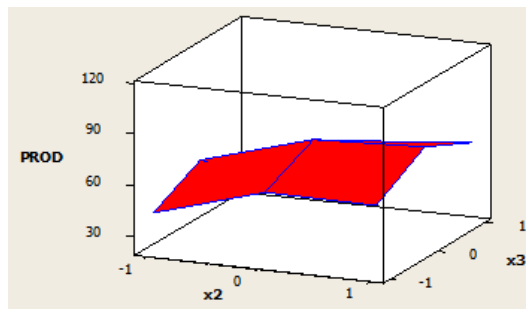


Figura 11.2.3.1. Gráfico considerando o fósforo (x2) e potássio (x3).

11.3. DELINEAMENTO COMPOSTO CENTRAL

O Delineamento Composto Central (DCC) ou Planejamento Composto Central (PCC) é um planejamento fatorial de primeira ordem, aumentado por pontos centrais e axiais, cujo objetivo é permitir a estimação dos parâmetros de uma superfície de resposta de segunda ordem.

Os pontos axiais são situados com qualquer distância $\pm \alpha$ da origem nos eixos do sistema de coordenadas. Os pontos axiais contribuem para a estimativa dos termos quadráticos. Os pontos fatoriais participam na estimativa dos termos lineares e interações entre fatores. Os pontos centrais, quando repetidos, informam sobre a existência de curvatura e estimam o erro puro. Caso seja observada a existência de curvatura, a adição de pontos axiais contribuirá de forma eficiente na obtenção das estimativas dos termos quadráticos.

11.3.1. Delineamento composto central para $2^k = 8$ pontos fatoriais

O DCC possui cinco níveis para um fatorial 2^k . Considerando $k = 3$ e um ponto central ($n_c = 1$), resultará em $N = 15$ pontos (tratamentos), sendo $n_f = 2^k = 8$ combinações fatoriais e $n_\alpha = 2k = 6$ pontos axiais.

$$N = n_f + n_\alpha + n_c = 2^k + 2k + 1 = 15$$

Tabela 11.3.1.1. Matriz de planejamento para um DCC ($N = n_f + n_\alpha + n_c = 15$).

	a_0	a_1	a_2	a_3	a_{11}	a_{22}	a_{33}	a_{12}	a_{13}	a_{23}
1	1	-1	-1	-1	$1-c$	$1-c$	$1-c$	1	1	1
2	1	-1	-1	1	$1-c$	$1-c$	$1-c$	1	-1	-1
3	1	-1	1	-1	$1-c$	$1-c$	$1-c$	-1	1	-1
4	1	-1	1	1	$1-c$	$1-c$	$1-c$	-1	-1	1
5	1	1	-1	-1	$1-c$	$1-c$	$1-c$	-1	-1	1
6	1	1	-1	1	$1-c$	$1-c$	$1-c$	-1	1	-1
7	1	1	1	-1	$1-c$	$1-c$	$1-c$	1	-1	-1
8	1	1	1	1	$1-c$	$1-c$	$1-c$	1	1	1
9	1	0	0	0	$-c$	$-c$	$-c$	0	0	0
10	1	α	0	0	$\alpha^2 - c$	$-c$	$-c$	0	0	0
11	1	$-\alpha$	0	0	$\alpha^2 - c$	$-c$	$-c$	0	0	0
12	1	0	α	0	$-c$	$\alpha^2 - c$	$-c$	0	0	0
13	1	0	$-\alpha$	0	$-c$	$\alpha^2 - c$	$-c$	0	0	0
14	1	0	0	α	$-c$	$-c$	$\alpha^2 - c$	0	0	0
15	1	0	0	$-\alpha$	$-c$	$-c$	$\alpha^2 - c$	0	0	0

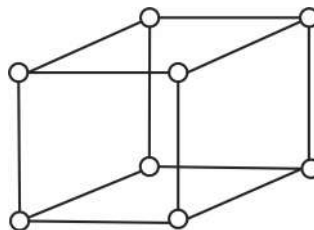
$$8(1-c) - 5c + 2(\alpha^2 - 1) = 0 \rightarrow c = \frac{2\alpha^2 + 8}{15}$$

$$c = \frac{2\alpha^2 + n_f}{n_f + n_\alpha + n_c}$$

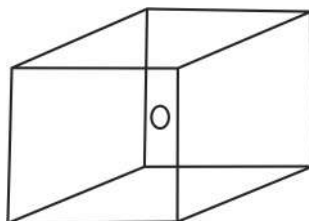
n_f : é o número dos pontos fatoriais;

$n_\alpha + n_c$: é o número de pontos axiais (estrelas) mais o número de pontos centrais.

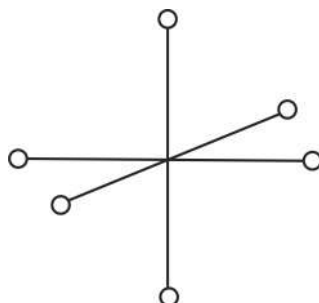
a) Pontos fatoriais ou porção cúbica ($n_f = 8$):



b) Ponto central ou porção central ($n_c = 1$):



c) Pontos axiais (estrelas) ou porção axial ($n_\alpha = 2k = 6$):



Considerando apenas um ponto central:

$$X'X = J \times \begin{bmatrix} 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 + 2\alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p & q & q & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q & p & q & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q & q & p & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix}$$

Em que:

$$p = \frac{n_f(n_\alpha + n_c) - 4n_f\alpha^2 - 4\alpha^4 + 2(n_f + n_\alpha + n_c)\alpha^4}{n_f + n_\alpha + n_c}$$

$$q = \frac{n_f(n_\alpha + n_c) - 4n_f\alpha^2 - 4\alpha^4}{n_f + n_\alpha + n_c}$$

$$\alpha = \left[\frac{(\sqrt{n_f + n_\alpha + n_c} - \sqrt{n_f})^2 \times n_f}{4} \right]^{\frac{1}{4}}$$

$$\alpha = \left[\frac{(\sqrt{n_f + n_\alpha + n_c} - \sqrt{n_f})^2 \times n_f}{4} \right]^{\frac{1}{4}} = \left[\frac{(\sqrt{15} - \sqrt{8})^2 \times 8}{4} \right]^{\frac{1}{4}} = 1,2154$$

Elemento (1,1): $N = n_f + n_\alpha + n_c = 8 + 6 + 1 = 15$;

Elementos diagonais (coeficientes de primeira ordem): $n_f + 2\alpha^2 = 8 + 2\alpha^2$;

Matriz simétrica de ordem k (coeficientes quadráticos): p e q ;

Elementos diagonais (coeficientes de interação): $n_f = 8$.

No caso de inserir apenas um ponto central, verifica-se que os termos quadráticos não são ortogonais. A condição para torná-los ortogonais é que o valor de q seja nulo:

$$n_f(n_\alpha + n_c) - 4n_f\alpha^2 - 4\alpha^4 = 0$$

Para obter-se a ortogonalidade dos termos quadráticos é necessário dimensionar o número dos pontos centrais. Considerando cinco pontos centrais ($n_c = 5$), resultará em $N = n_f + n_\alpha + n_c = 8 + 6 + 5 = 19$ pontos no total. Logo:

$$\alpha = \left[\frac{(\sqrt{n_f + n_\alpha + n_c} - \sqrt{n_f})^2 n_f}{4} \right]^{\frac{1}{4}} = \left[\frac{(\sqrt{19} - \sqrt{8})^2 \times 8}{4} \right]^{\frac{1}{4}} = 1,47119475$$

Para $q = 0$, implica que $n_f(n_\alpha + n_c) - 4n_f\alpha^2 - 4\alpha^4 = 0$, então:

$$n_f(n_\alpha + n_c) - 4n_f(1,47119475)^2 - 4(1,47119475)^4 = 0$$

$$8 \times 1 - 4 \times 8 \times 1,47119475^2 - 4 \times 1,47119475^4 = 0$$

O DCC com $2^k = 2^3 = 8$ pontos fatoriais e cinco pontos centrais com $\alpha = 1,47119475$ é um delineamento ortogonal.

No caso do DCC com $2^k = 2^2 = 4$ pontos fatoriais, $n_\alpha = 2k = 4$ (quatro pontos axiais) e $n_c = 5$, então: $N = n_f + n_\alpha + n_c = 4 + 4 + 5 = 13$. Logo:

$$\alpha = \left[\frac{(\sqrt{n_f + n_\alpha + n_c} - \sqrt{n_f})^2 n_f}{4} \right]^{\frac{1}{4}} = \left[\frac{(\sqrt{13} - \sqrt{4})^2 \times 4}{4} \right]^{\frac{1}{4}} = 1,26710350$$

$$n_f(n_\alpha + n_c) - 4n_f\alpha^2 - 4\alpha^4 = 0$$

$$n_f(n_\alpha + n_c) - 4n_f(1,26710350)^2 - 4(1,26710350)^4 = 0$$

$$4 \times 9 - 4 \times 4 \times 1,26710350^2 - 4 \times 1,26710350^4 = 0$$

O DCC com $2^k = 2^2 = 4$ pontos fatoriais e cinco pontos centrais com $\alpha = 1,26710350$ é um delineamento ortogonal.

De acordo com Mendonça (2012), para um determinado número de fatores, dependendo dos valores assumidos por α e pelo número de repetições do ponto central, o DCC possibilita ao pesquisador escolher entre os tipos ortogonais e/ou rotacionais e, ainda, a sua divisão ortogonalmente em blocos, bem como trabalhar, em relação aos fatoriais completos, com um número menor de combinações entre os níveis dos diversos fatores estudados.

A escolha do valor de α é muito importante no planejamento de um experimento. Qualquer valor pode ser dado para α . Contudo, é o valor α que caracteriza o DCC, pois ele pode ser definido para tornar os coeficientes da regressão ortogonais, diminuir o desvio da regressão da superfície de resposta à sua forma verdadeira, ou para dar ao delineamento a propriedade de ser rotacional. No DCC esférico, o valor de alfa é igual a $\alpha = \sqrt{k}$, o que não seria a melhor opção. O DCC rotacional garante que a variância $Var(\hat{Y})$ do modelo de previsão seja a mesma para qualquer raio constante, em que o valor de α é igual a $\alpha = \sqrt[4]{2^k}$. Na prática, o DCC rotacional é o mais utilizado.

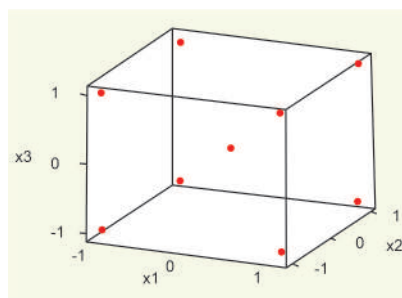
É importante observar, que quando $k = 2$, ocorre que $\alpha = \sqrt[4]{2^k} = \sqrt[4]{2^2}$. Isto é, $\alpha = \sqrt{k} = \sqrt{2}$, ou seja, o DCC para $k = 2$ é rotacional e também esférico.

No caso de $k = 3$, ocorre o valor $\alpha = \sqrt{k} = \sqrt{3} = 1,7321$ para o DCC esférico e para o DCC rotacional $\alpha = \sqrt[4]{2^k} = \sqrt[4]{2^3} = \sqrt[4]{8} = 1,6818$.

O uso de experimentos fatoriais com dois níveis com adição de pontos centrais é essencial no planejamento de superfície de resposta, na pesquisa de estratégia sequencial para a experimentação e na investigação da região de ótimo, quer dizer, na busca de uma modelagem passível de otimização. Tais princípios devem ser o foco da metodologia de superfície de resposta, em que os objetivos são: testar a curvatura e estimar o erro experimental.

A utilização de pontos centrais promove a redução de recursos na experimentação, pois replicar no fatorial tende a aumentar os custos. A melhor alternativa é trabalhar o fatorial 2^k , adicionando algumas replicações codificadas com zero no ponto central, o que significa uma condição intermediária.

Supondo o uso de fatores quantitativos, depreendendo uma escala contínua, seja considerado o exemplo de um fatorial 2^3 com pontos centrais, em que o número de combinações seria o número de pontos fatoriais mais o número de pontos centrais, ou seja, $N = 2^k + n_c$. O ponto central, além de ser responsável também pela estimativa dos efeitos quadráticos, possibilita a estimativa do erro puro, quando ele é repetido.



Assume-se, no planejamento de um fatorial 2^k , a premissa de que os fatores apresentem efeitos lineares. Entretanto, com adição de pontos centrais, pode-se testar se há necessidade de usar um modelo quadrático, o qual pode possibilitar um melhor ajuste em comparação com o modelo somente baseado nos efeitos principais e nas suas respectivas interações.

A adição de pontos centrais no fatorial 2^k possibilita, mas não garante, a obtenção dos coeficientes quadráticos das variáveis, pois apenas testa se há a necessidade de suas inclusões, por meio do teste de curvatura no modelo de regressão.

Seja considerado o modelo para o fatorial, mas somente com os termos de segunda ordem mistos, isto é, as interações de segunda ordem:

$$Y = a_0 + \sum_{j=1}^k a_j x_j + \sum \sum a_{ij} x_i x_j + e$$

O ideal é testar um modelo quadrático mais completo com a inclusão dos termos quadráticos:

$$Y_i = a_0 + \sum_{j=1}^k a_j x_j + \sum_{i=1}^k \sum_{j=1}^k a_{ij} x_i x_j + \sum_{j=1}^k a_{jj} x_{jj}^2 + e_i \quad i < j$$

O fatorial com pontos centrais possibilita testar a inclusão dos termos quadráticos pelo teste de curvatura:

Sejam consideradas as hipóteses a serem utilizadas no teste de curvatura:

$$H_0 : \sum_{j=1}^k a_{jj} = 0 \text{ vs } H_A : \sum_{j=1}^k a_{jj} \neq 0$$

No caso da rejeição da hipótese H_0 , é importante completar o planejamento e viabilizar a inclusão dos coeficientes quadráticos. Entretanto, como no fatorial 2^k somente com pontos centrais não é possível obtê-los, pois no mínimo teria que haver combinações dos pontos centrais com outros níveis mais distantes, surgindo a necessidade de inserir os pontos axiais.

A alocação de pontos centrais é importante para produzir um planejamento experimental robusto, não suscetível ao aparecimento de *outliers*, assim como para minimizar a variância experimental, permitindo verificar a falta de ajuste. De acordo com Carpinetti (2009) é possível testar estatisticamente a existência ou não de curvatura na região central da superfície de resposta, por meio do teste “ $F_{1; n_c-1; \alpha}$ ”, obtendo-se a soma de quadrados da curvatura pela expressão:

$$SQ_{CURVATURA} = \frac{n_f \times n_c (\bar{y}_f - \bar{y}_c)^2}{n_f + n_c}$$

$\bar{y}_f$: valor médio estimado dos valores dos pontos fatoriais;

$\bar{y}_c$: valor médio estimado dos valores dos pontos centrais.

Sendo:

$$F = \frac{SQ_{CURVATURA}}{\hat{\sigma}^2}$$

$\hat{\sigma}^2$ é a estimativa do erro experimental obtida a partir dos valores dos pontos centrais repetidos e alocados no ponto central.

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n_c} (y_{ci} - \bar{y}_c)^2}{n_c - 1}$$

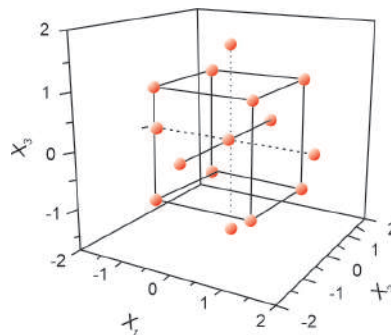
Exercício 11.3.1.1. Seja um ensaio de adubação NPK de uma espécie florestal, considerando o delineamento composto central inteiramente ao acaso com seis pontos axiais, tal que $\alpha = 2$ e um ponto central, totalizando 15 tratamentos com duas repetições ($J = 2$). Na tabela 11.3.1.2, estão os dados referentes à produção de madeira em m^3 por parcela, considerando as doses 0, 60, 120, 180 e 240 kg por parcela.

Tabela 11.3.1.2. Produção de madeira (Dados simulados).

Trat.	Rep1	Rep2	Total	Trat.	Rep1	Rep2	Total
111	19,1	25,9	45,0	222	22,8	30,8	53,6
113	15,4	18,5	33,9	422	12,3	18,5	30,8
131	24,0	29,6	53,6	022	20,9	12,9	33,8
133	32,7	21,3	54,0	242	12,9	16,0	28,9
311	29,0	18,5	47,5	202	14,1	12,3	26,4
313	25,3	29,6	54,9	224	17,8	16,0	33,8
331	27,7	29,6	57,3	220	18,5	17,8	36,3
333	32,0	29,6	61,6				

Tal que 222 é o ponto central, assim como:

	0	1	2	3	4
x_i	-2	-1	0	1	2



Seja considerado o modelo de segunda ordem:

$$Y_i = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{11}(x_1^2 - c) + a_{22}(x_2^2 - c) + a_{33}(x_3^2 - c) + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + e_i$$

$$c = \frac{2\alpha^2 + n_f}{n_f + n_\alpha + n_c} = \frac{2 \times 2^2 + 8}{8 + 6 + 1} = \frac{16}{15}$$

Tabela 11.3.1.3. Matriz X de planejamento da regressão polinomial de segunda ordem.

a_0	a_1	a_2	a_3	a_{11}	a_{22}	a_{33}	a_{12}	a_{13}	a_{23}
1	-1	-1	-1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	1	1	1
1	-1	-1	1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	1	-1	-1
1	-1	1	-1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	-1	1	-1
1	-1	1	1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	-1	-1	1
1	1	-1	-1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	-1	-1	1
1	1	-1	1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	-1	1	-1
1	1	1	-1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	1	-1	-1
1	1	1	1	$-\frac{1}{15}$	$-\frac{1}{15}$	$-\frac{1}{15}$	1	1	1
1	0	0	0	$-\frac{16}{15}$	$-\frac{16}{15}$	$-\frac{16}{15}$	0	0	0
1	2	0	0	$\frac{44}{15}$	$-\frac{16}{15}$	$-\frac{16}{15}$	0	0	0
1	-2	0	0	$\frac{44}{15}$	$-\frac{16}{15}$	$-\frac{16}{15}$	0	0	0
1	0	2	0	$-\frac{16}{15}$	$\frac{44}{15}$	$-\frac{16}{15}$	0	0	0
1	0	-2	0	$-\frac{16}{15}$	$\frac{44}{15}$	$-\frac{16}{15}$	0	0	0
1	0	0	2	$-\frac{16}{15}$	$-\frac{16}{15}$	$\frac{44}{15}$	0	0	0
1	0	0	-2	$-\frac{16}{15}$	$-\frac{16}{15}$	$\frac{44}{15}$	0	0	0

$$X'X = 2 \times \begin{bmatrix} 15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 16 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{344}{15} & -\frac{136}{15} & -\frac{136}{15} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{136}{15} & \frac{344}{15} & -\frac{136}{15} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{136}{15} & -\frac{136}{15} & \frac{344}{15} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix}$$

Em que:

$$p = \frac{n_f(n_\alpha + n_c) - 4n_f\alpha^2 - 4\alpha^4 + 2(n_f + n_\alpha + n_c)\alpha^4}{n_f + n_\alpha + n_c} = \frac{344}{15}$$

$$q = \frac{n_f(n_\alpha + n_c) - 4n_f\alpha^2 - 4\alpha^4}{n_f + n_\alpha + n_c} = \frac{-136}{15}$$

a) Estimativas dos parâmetros

Dado que $\hat{\beta}_v = (X'X)^{-1} X'Y_v$ em que $J = 2$, então:

$$X'Y_v = \begin{bmatrix} 651,4000 \\ 28,8000 \\ 50,2000 \\ -4,0000 \\ -28,6267 \\ -65,8267 \\ -6,6267 \\ -12,2000 \\ 22,4000 \\ 8,400 \end{bmatrix} \quad \hat{\beta}_v = \frac{1}{2}(X'X)^{-1}(X'Y_v) = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_{11} \\ \hat{a}_{22} \\ \hat{a}_{33} \\ \hat{a}_{12} \\ \hat{a}_{13} \\ \hat{a}_{23} \end{bmatrix} = \begin{bmatrix} 21,7133 \\ 0,9000 \\ 1,5688 \\ -0,1250 \\ -3,4306 \\ -4,0118 \\ -3,0868 \\ -0,7625 \\ 1,4000 \\ 0,5250 \end{bmatrix}$$

Em que: $Y_v = [45,0 \ 33,9 \ 53,6 \ 54,0 \ 47,5 \ 54,9 \ 57,3 \ 61,6 \ 53,6 \ 30,8 \ 33,8 \ 28,9 \ 26,4 \ 33,8 \ 36,3]$

Tabela 11.3.1.4. Estimativas dos parâmetros obtidos no SAS.

Parâmetros	G.L.	Estimativas	Erro Padrão	Valor "t"	Pr > t	$x_* = \frac{x}{2}$
Intercepto	1	32,944444	3.756973	8,77	< 0,0001	32,944444
N	1	0,900000	1.065002	0,85	0,4081	1,800000
P	1	1,568750	1.065002	1,47	0,1563	3,137500
K	1	-0,125000	1.065002	-0,12	0,9077	-0,250000
N×N	1	-3,430556	1.279973	-2,68	0,0144	-13,722222
P×N	1	-0,762500	1.506140	-0,51	0,6182	-3,050000
P×P	1	-4,011806	1.279973	-3,13	0,0052	-16,047222
K×N	1	1,400000	1.506140	0,93	0,3637	5,600000
K×P	1	0,525000	1.506140	0,35	0,7311	2,100000
K×K	1	-3,086806	1.279973	-2,41	0,0256	-12,347222

Considerando $c = \frac{16}{15}$, logo:

$$\hat{Y}_i = 21,7133 + 0,9x_1 + 1,5688x_2 - 0,125x_3 - 3,4306(x_1^2 - c) - 4,0118(x_2^2 - c) - 3,0868(x_3^2 - c) - 0,7625x_1x_2 + 1,4x_1x_3 + 0,525x_2x_3$$

$$\hat{Y}_i = 32,9444 + 0,9x_1 + 1,5688x_2 - 0,125x_3 - 3,4306x_1^2 - 4,0118x_2^2 - 3,0868x_3^2 - 0,7625x_1x_2 + 1,4x_1x_3 + 0,525x_2x_3 \quad (11.3.1.1)$$

Na tabela 11.3.1.4 estão apresentadas as estimativas dos parâmetros obtidos no SAS, conforme o programa **Prog36**, encontrado no apêndice B em anexo.

$$\hat{Y}_i = 32,9444 + 1,8x_{*1} + 3,1375x_{*2} - 0,25x_{*3} - 13,72222x_{*1}^2 - 16,047222x_{*2}^2 - 12,347222x_{*3}^2 - 3,05x_{*1}x_{*2} + 5,6x_{*1}x_{*3} + 2,1x_{*2}x_{*3} \quad 11.3.1.2$$

Esses valores diferentes no SAS resultam de uma codificação para os coeficientes das variáveis independentes, as quais são divididas por dois, isto é, $x_* = x/2$, em que $x = 2x_*$. Denota-se que os coeficientes das variáveis independentes da equação 11.3.1.2 correspondem ao dobro dos coeficientes da equação 11.3.1.1.

b) Dose economicamente aconselhável

Dado o modelo de superfície de resposta:

$$Y_i = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^k a_{ii} x_i^2 + \sum_i \sum_j a_{ij} x_i x_j + e_i \quad i < j$$

Em que:

$$\sum_{i=1}^k a_i x_i : \text{componentes de primeira ordem;}$$

$$\sum_{i=1}^k \sum_{i=1}^k a_{ii} x_i^2 + \sum_i \sum_j a_{ij} x_i x_j : \text{componentes de segunda ordem.}$$

Considerando um ensaio de adubação de uma espécie florestal, tal que RL seja a equação da receita líquida, então:

$$RL = w\hat{Y}_i - cf - q_1 t_1 x_1 - q_2 t_2 x_2 - q_3 t_3 x_3$$

$$\frac{RL}{w} = \hat{Y}_i - \frac{cf}{w} - \frac{t_1 q_1 x_1}{w} - \frac{t_2 q_2 x_2}{w} - \frac{t_3 q_3 x_3}{w}$$

$$\frac{dRL}{w} = d\hat{Y}_i - \frac{t_1}{w} q_1 dx_1 - \frac{t_2}{w} q_2 dx_2 - \frac{t_3}{w} q_3 dx_3$$

$$\begin{cases} 2\hat{a}_{11}x_1^* + \hat{a}_{12}x_2^* + \hat{a}_{13}x_3^* = \frac{t_1q_1}{w} - \hat{a}_1 \\ \hat{a}_{12}x_1^* + 2\hat{a}_{22}x_2^* + \hat{a}_{23}x_3^* = \frac{t_2q_2}{w} - \hat{a}_2 \\ \hat{a}_{13}x_1^* + \hat{a}_{23}x_2^* + 2\hat{a}_{33}x_3^* = \frac{t_3q_3}{w} - \hat{a}_3 \end{cases}$$

Seja: $q_1 = q_2 = q_3 = 60$ e $\frac{t_1}{w} = \frac{t_2}{w} = \frac{t_3}{w} = 0,10$, então:

$$\begin{cases} 2 \times -3,4306x_1^* - 0,7625x_2^* + 1,4x_3^* = 6 - 0,9 = 5,1 \\ -0,7625x_1^* + 2 \times -4,0118x_2^* + 0,525x_3^* = 6 - 1,5688 = 4,4312 \\ 1,4x_1^* + 0,525x_2^* + 2 \times -3,0868x_3^* = 6 + 0,125 = 6,125 \end{cases}$$

$$\begin{cases} -6,8612x_1^* - 0,7625x_2^* + 1,4x_3^* = 5,1 \\ -0,7625x_1^* - 8,0236x_2^* + 0,525x_3^* = 4,4312 \\ 1,4x_1^* + 0,525x_2^* - 6,1736x_3^* = 6,125 \end{cases}$$

$$x^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} -6,8612 & -0,7625 & 1,4 \\ -0,7625 & -8,0236 & 0,525 \\ 1,4 & 0,525 & -6,1736 \end{bmatrix}^{-1} \begin{bmatrix} 5,1 \\ 4,4312 \\ 6,125 \end{bmatrix}$$

$$x^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} -0,15404 & 0,0124222 & -0,033876 \\ 0,0124222 & -0,126331 & -0,007926 \\ -0,033876 & -0,007926 & -0,170336 \end{bmatrix} \begin{bmatrix} 5,1 \\ 4,4312 \\ 6,125 \end{bmatrix}$$

$$x_v^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} -0,938045 \\ -0,544995 \\ -1,251196 \end{bmatrix}$$

Dado que:

X_i	0	60	120	180	240
x_i	-2	-1	0	1	2

$$-0,938045 = \frac{X_1^* - 120}{60} \quad -0,544995 = \frac{X_2^* - 120}{60} \quad -1,251196 = \frac{X_3^* - 120}{60}$$

Logo:

$$X_v^* = \begin{bmatrix} X_1^* \\ X_2^* \\ X_3^* \end{bmatrix} = \begin{bmatrix} 63,717281 \\ 87,300325 \\ 44,928221 \end{bmatrix}$$

Considerando que:

$$\hat{Y}_i = 32,9444 + 0,9x_1 + 1,5688x_2 - 0,125x_3 - 3,4306x_1^2 - 4,0118x_2^2 - 3,0868x_3^2 - 0,7625x_1x_2 + 1,4x_1x_3 + 0,525x_2x_3$$

$$x_v^* = \begin{bmatrix} x_1^* \\ x_2^* \\ x_3^* \end{bmatrix} = \begin{bmatrix} -0,938045 \\ -0,544995 \\ -1,251196 \end{bmatrix}$$

Definindo $t_1/w = t_2/w = t_3/w = 0,10$, o valor da produção de madeira para x_v^* é dado por:

$$\hat{Y}^* = 23,9703 \text{ m}^3 \text{ por parcela}$$

c) Obtenção do ponto de máximo sem considerar o aspecto econômico:

$$x_{v0} = \begin{bmatrix} x_{01} \\ x_{02} \\ x_{03} \end{bmatrix} = \begin{bmatrix} -6,8612 & -0,7625 & 1,4 \\ -0,7625 & -8,0236 & 0,525 \\ 1,4 & 0,525 & -6,1736 \end{bmatrix}^{-1} \begin{bmatrix} -0,9 \\ -1,5688 \\ 0,125 \end{bmatrix}$$

$$x_{v0} = \begin{bmatrix} x_{01} \\ x_{02} \\ x_{03} \end{bmatrix} = \begin{bmatrix} 0,1149134 \\ 0,1860181 \\ 0,0216305 \end{bmatrix}$$

Alternativamente, pode-se obter o ponto de máximo da seguinte forma:

$$\hat{Y}_i = 32,9444 + 0,9x_1 + 1,5688x_2 - 0,125x_3 - 3,4306x_1^2 - 4,0118x_2^2 - 3,0868x_3^2 - 0,7625x_1x_2 + 1,4x_1x_3 + 0,525x_2x_3$$

$$x_{v0} = -\frac{1}{2}A^{-1}\hat{a}_v = -\frac{1}{2} \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12}/2 & \hat{a}_{13}/2 \\ \hat{a}_{12}/2 & \hat{a}_{22} & \hat{a}_{23}/2 \\ \hat{a}_{13}/2 & \hat{a}_{23}/2 & \hat{a}_{33} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{bmatrix}$$

$$x_{v0} = \begin{bmatrix} -3,4306 & -0,38125 & 0,7 \\ -0,38125 & -4,0118 & 0,2625 \\ 0,7 & 0,2625 & -3,0868 \end{bmatrix}^{-1} \begin{bmatrix} 0,9 \\ 1,5688 \\ -0,125 \end{bmatrix} = \begin{bmatrix} 0,1149134 \\ 0,1860181 \\ 0,0216305 \end{bmatrix}$$

Dado que:

$$0,1149134 = \frac{X_{01} - 120}{60} \quad 0,1860181 = \frac{X_{02} - 120}{60} \quad 0,0216305 = \frac{X_{03} - 120}{60}$$

Logo:

$$X_{v0} = \begin{bmatrix} X_{01} \\ X_{02} \\ X_{03} \end{bmatrix} = \begin{bmatrix} 126,89481 \\ 131,16108 \\ 121,29783 \end{bmatrix}$$

d) Análise da natureza do ponto crítico

Dada a matriz Hessiana A :

$$A = \begin{bmatrix} -3,4306 & -0,38125 & 0,7 \\ -0,38125 & -4,0118 & 0,2625 \\ 0,7 & 0,2625 & -3,0868 \end{bmatrix}$$

Então:

$$\det(A - \lambda I) = \left\{ \begin{bmatrix} -3,4306 & -0,38125 & 0,7 \\ -0,38125 & -4,0118 & 0,2625 \\ 0,7 & 0,2625 & -3,0868 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} = 0$$

Resultando em:

$$\hat{\lambda} = \begin{bmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3 \end{bmatrix} = \begin{bmatrix} -2,537282 \\ -3,533775 \\ -4,458143 \end{bmatrix}$$

Como os três autovalores são negativos, então $X_v = (X_{01} \ X_{02} \ X_{03})$ é um ponto estacionário de máximo, logo a superfície de resposta tem convexidade para baixo. Seja M a matriz dos autovetores normalizados associados aos autovalores obtidos:

$$M = \begin{bmatrix} 0,622360 & -0,553864 & 0,553084 \\ -0,02163 & 0,694172 & 0,719484 \\ 0,782432 & 0,459739 & -0,420048 \end{bmatrix}$$

O programa **Prog37**, em anexo no apêndice B, efetua as operações matriciais usadas nos cálculos. No caso da dosagem economicamente aconselhável, em que a matriz A é definida negativa, conclui-se para razão $t_1/w = t_2/w = t_3/w = 0,10$ que o ponto crítico x_v^* ($N = 63,717281$ $P = 87,300325$ $K = 44,92822$) é um ponto de máximo.

Não considerando o aspecto econômico e substituindo o ponto de máximo $x_{v0} = [0,1149134 \ 0,1860181 \ 0,02116305]$ na equação:

$$\hat{Y}_i = 32,9444 + 0,9x_1 + 1,5688x_2 - 0,125x_3 - 3,4306x_1^2 - 4,0118x_2^2 - 3,0868x_3^2 - 0,7625x_1x_2 + 1,4x_1x_3 + 0,525x_2x_3$$

Logo:

$$\hat{Y}_{Max} = \hat{Y}(0,1149134, 0,1860181, 0,0263305) = 33,1407 \text{ m}^3 \text{ por parcela.}$$

É importante observar que os pontos de máximos obtidos, com ou sem condicionante econômica, estão dentro da região experimental, pois foram usadas as doses 0, 60, 120, 180 e 240 kg por parcela.

x_i	-2	-1	0	1	2
-------	----	----	---	---	---

Da obtenção de ponto de máximo, sem considerar o aspecto econômico:

$$\hat{\lambda} = \begin{bmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3 \end{bmatrix} = \begin{bmatrix} -2,537282 \\ -3,533775 \\ -4,458143 \end{bmatrix}$$

Logo, seja considerada a equação canônica:

$$\hat{Y} = 33,140707 - 2,537282\hat{\lambda}_1^2 - 3,533775\hat{\lambda}_2^2 - 4,458143\hat{\lambda}_3^2$$

Tabela 11.3.1.5. Valores críticos, autovalores e autovetores (SAS).

Fatores	Para $x_* = x/2$	Para $x = 2x_*$	Autovalores	Autovetores		
				N	P	K
N	0,057458	0,114916	-10,149072	0,622360	-0,02163	0,782432
P	0,093006	0,186011	-14,135061	-0,553864	0,694172	0,459739
K	0,010815	0,021630	-17,832534	0,553084	0,719484	-0,420048
Valor estimado do ponto estacionário: 33,140707			O ponto estacionário é um máximo			

Conforme o programa SAS, **Prog36**, em anexo no apêndice B, a tabela 11.3.1.5 apresenta os resultados sobre a natureza do ponto estacionário. A figura 11.3.1.1 mostra a superfície, considerando o fósforo e o potássio que apresentam os maiores valores ($|\lambda_2|$ e $|\lambda_3|$) dos autovalores, isto é, $\hat{\lambda}_2 = -3,533775$ e $\hat{\lambda}_3 = -4,458143$. A produção máxima calculada corresponde a $\hat{Y}_0 = 33,140707 \text{ m}^3$ por parcela.

No caso da equação:

$$\begin{aligned}\hat{Y} = & 32,9444 + 1,8x_{*1} + 3,1375x_{*2} - 0,25x_{*3} - 13,72222x_{*1}^2 - 16,047222x_{*2}^2 \\ & - 12,347222x_{*3}^2 - 3,05x_{*1}x_{*2} + 5,6x_{*1}x_{*3} + 2,1x_{*2}x_{*3}\end{aligned}$$

Sendo o ponto estacionário $x'_{v0} = [0,057458 \ 0,093006 \ 0,010815]$ e os autovalores dados por: $\hat{\lambda}'_* = [-10,149072 \ -14,135061 \ -17,832534]$, as conclusões são as mesmas.

$$x_{v0} = -\frac{1}{2}A^{-1}\hat{a}_v = -\frac{1}{2}\begin{bmatrix}\hat{a}_{11} & \hat{a}_{12}/2 & \hat{a}_{13}/2 \\ \hat{a}_{12}/2 & \hat{a}_{22} & \hat{a}_{23}/2 \\ \hat{a}_{13}/2 & \hat{a}_{23}/2 & \hat{a}_{33}\end{bmatrix}^{-1}\begin{bmatrix}\hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3\end{bmatrix}$$

$$x_{v0} = -\frac{1}{2}\begin{bmatrix}-13,72222 & -1,525 & 2,8 \\ -1,525 & -16,047222 & 1,05 \\ 2,8 & 1,05 & -12,347222\end{bmatrix}^{-1}\begin{bmatrix}1,8 \\ 3,1375 \\ -0,25\end{bmatrix}$$

$$x_{v0} = \begin{bmatrix}0,057458 \\ 0,093006 \\ 0,010815\end{bmatrix}$$

$$A = \begin{bmatrix}-13,72222 & -1,525 & 2,8 \\ -1,525 & -16,047222 & 1,05 \\ 2,8 & 1,05 & -12,347222\end{bmatrix}$$

$$\det(A - \lambda I) = \left\{ A = \begin{bmatrix}-13,72222 & -1,525 & 2,8 \\ -1,525 & -16,047222 & 1,05 \\ 2,8 & 1,05 & -12,347222\end{bmatrix} - \lambda \begin{bmatrix}1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{bmatrix} \right\} = 0$$

$$\lambda'_* = \begin{bmatrix}\lambda'_{*1} \\ \lambda'_{*2} \\ \lambda'_{*3}\end{bmatrix} = \begin{bmatrix}-10,149072 \\ -14,135061 \\ -17,832534\end{bmatrix}$$

Em que M é a matriz dos autovetores normalizados:

$$M = \begin{bmatrix}0,622360 & -0,553864 & 0,553084 \\ -0,02163 & 0,694172 & 0,719484 \\ 0,782432 & 0,459739 & -0,420048\end{bmatrix}$$

$$\text{Observar que } \hat{\lambda}_* = \begin{bmatrix}\hat{\lambda}_{*1} \\ \hat{\lambda}_{*2} \\ \hat{\lambda}_{*3}\end{bmatrix} = 4 \times \begin{bmatrix}\hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3\end{bmatrix} = 4 \times \begin{bmatrix}-2,537282 \\ -3,533775 \\ -4,458143\end{bmatrix} = \begin{bmatrix}-10,1491 \\ -14,1351 \\ -17,8325\end{bmatrix}$$

Em que:

$$\hat{\lambda} = \begin{bmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3 \end{bmatrix} = \begin{bmatrix} -2,537282 \\ -3,533775 \\ -4,458143 \end{bmatrix}$$

São os autovalores referentes à equação:

$$\hat{Y}_i = 32,9444 + 0,9x_1 + 1,5688x_2 - 0,125x_3 - 3,4306x_1^2 - 4,0118x_2^2 - 3,0868x_3^2 - 0,7625x_1x_2 + 1,4x_1x_3 + 0,525x_2x_3$$

Seja considerada a matriz M dos autovetores normalizados associados ao vetor $\hat{\lambda}$:

$$M = \begin{bmatrix} 0,622360 & -0,553864 & 0,553084 \\ -0,02163 & 0,694172 & 0,719484 \\ 0,782432 & 0,459739 & -0,420048 \end{bmatrix}$$

Observação: as matrizes dos autovetores normalizados são as mesmas para ambas as equações:

Dado que a matriz M é ortogonal, logo $M^{-1} = M'$ e então $MM' = I$. Seja $u_v = M'z_v$, em que $z_v = x_v - x_{v0}$.

$$x_{v0} = \begin{bmatrix} 0,1149134 \\ 0,1860181 \\ 0,026305 \end{bmatrix} \quad z_v = x_v - x_{v0} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} 0,1149134 \\ 0,1860181 \\ 0,026305 \end{bmatrix}$$

A relação entre as variáveis $x'_v = [x_1, x_2, x_3]$ e $u'_v = [u_1, u_2, u_3]$:

$$u_v = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = M'z_v = \begin{bmatrix} 0,622360 & -0,02163 & 0,782432 \\ -0,553864 & 0,694172 & 0,459739 \\ 0,553084 & 0,719484 & -0,420048 \end{bmatrix} \times \begin{bmatrix} x_1 - 0,1149134 \\ x_2 - 0,1860181 \\ x_3 - 0,026305 \end{bmatrix}$$

Quando ocorre ponto de sela ou o ponto crítico fora da região experimental, a relação entre as variáveis x_v e u_v é importante para estudar as regiões de interesse. Entretanto, nesse caso, o ponto estacionário é de máximo e ocorreu dentro da região experimental permitindo considerar a presente análise finalizada.

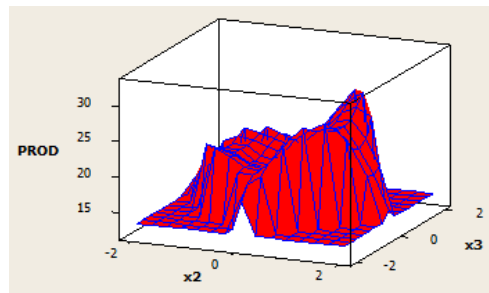


Figura 11.3.1.1. Gráfico considerando o fósforo e o potássio.

Por outro lado, considerando que o fósforo (x_2) e o potássio (x_3) apresentaram os maiores autovalores (valores absolutos), a figura 11.3.1.1 mostra a tendência do máximo em direção à produção de $\hat{Y}_0 = 33,140707 \text{ m}^3$ por parcela.

11.4. PLANEJAMENTO BOX-BEHNKEN

O Box-Behnken (1960), assim como o plano composto central, utiliza a metodologia de superfície de resposta, gerando equações polinomiais quadráticas.

No caso de três fatores, observa-se que, considerando x_1 e x_2 variando e fixando $x_3 = 0$, ocorre um fatorial 2^2 ; assim como quando x_1 e x_3 variam com o valor fixado $x_2 = 0$, resulta em um fatorial 2^2 ; o mesmo ocorrendo quando x_2 e x_3 variam com $x_1 = 0$ fixo. Na tabela 11.4.1, considerando três fatores, ocorre a seguinte a matriz de planejamento:

Tabela 11.4.1. Matriz de planejamento Box-Behnken para três fatores.

	x_1	x_2	x_3	
1	-1	-1	0	Fatorial 2^2 para x_1 e x_2 com $x_3 = 0$
2	+1	-1	0	
3	-1	+1	0	
4	+1	+1	0	
5	-1	0	-1	Fatorial 2^2 para x_1 e x_3 com $x_2 = 0$
6	+1	0	-1	
7	-1	0	+1	
8	+1	0	+1	
9	0	-1	-1	Fatorial 2^2 para x_2 e x_3 com $x_1 = 0$
10	0	+1	-1	
11	0	-1	+1	
12	0	+1	+1	
13	0	0	0	Pontos centrais
14	0	0	0	
15	0	0	0	

No delineamento Box-Behnken, exemplificando com $k = 3$, os pontos fatoriais não estão localizados nos vértices de um cubo, como no composto central, mas na metade das arestas. A figura 11.4.1 mostra as configurações dos delineamentos considerando, o fatorial completo, o composto central (DCC) com um ponto central e o Box-Behnken (DBB) com um ponto central. Por outro lado, observando a formatação do DBB, verifica-se que, para cada uma das 12 arestas, há um ponto experimental, além de que, quando considerado um ponto central, o total de combinações será igual a 13.

Em termos de propriedades, o planejamento Box-Behnken é econômico e eficiente. É rotacionável para $k = 4$ e $k = 7$, mas para k diferente de 4 e de 7 não perde muito em termos de rotacionalidade. Os delineamentos Box-Behnken são esféricos, ou seja, todos os pontos estão a uma mesma distância de $\sqrt{2}$ do centro do planejamento. Lima (2009) mostra, na figura 11.4.1, para melhor iden-

tificar as simetrias, duas representações do planejamento Box-Behnken para três fatores. Silva (2014) apresenta na figura 11.4.2 as representações do fatorial completo, DCC e DBB.

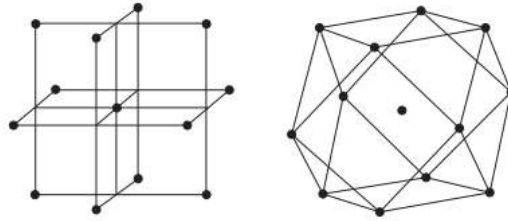


Figura 11.4.1. Duas representações do planejamento Box-Behnken para três fatores.

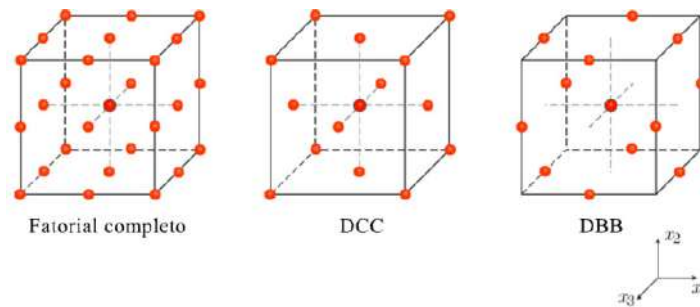


Figura 11.4.2. Representações do fatorial completo, DCC e DBB.

Em comparação com um planejamento fatorial completo de 3 fatores e 3 níveis, totalizando 27 combinações, o delineamento Box-Behnken com apenas 13 combinações, em que foi adicionado apenas um ponto central, mostra ser mais econômico por reduzir o número de combinações e ser mais eficiente. Também gera menos coeficientes no ajuste do modelo, uma vez que a eficiência de uma estrutura experimental é avaliada pela razão entre o número de coeficientes gerados e o número de combinações testadas.

Comparado a um planejamento composto central, um plano Box-Behnken com 3 variáveis (fatores) gera 13 pontos experimentais e um total de 10 coeficientes. Embora o composto central gere o mesmo número de coeficientes, apresenta 15 pontos experimentais, reduzindo, assim, a eficiência relativa do sistema, quando aplicada a razão coeficientes/número de combinações.

O planejamento Box-Behnken não é cúbico, pois não possui pontos nos vértices. Tal como alguns DCCs, é um planejamento esférico no sentido da distância, que será a mesma dos pontos ao ponto central.

O DCC tem pontos nos vértices, possuindo níveis nos pontos extremos, em que o raio do planejamento está para fora do cubo, assim como há um raio maior do que o Box-Behnken, o qual possui a restrição de só ser utilizado para $k \geq 3$, enquanto o DCC pode ser utilizado para $k = 2$.

O número de combinações no Box-Behnken será: $N = 2^2 \times \binom{k}{2} + n_c$, em que:

$$N = 2^2 \times \frac{k(k-1)}{2} + n_c = 2k(k-1) + n_c$$

O quadro abaixo apresenta, sem considerar os pontos centrais (n_c), a comparação entre o Box-Behnken (DBB) e o DCC, com relação ao número de combinações requeridas:

k	DCC	DBB
2	$8 + n_c$	-----
3	$14 + n_c$	$12 + n_c$
4	$24 + n_c$	$24 + n_c$
5	$26 + n_c$	$40 + n_c$

- a) Quando $k = 3$, o delineamento Box-Behnken [$2k(k - 1) = 12$] é mais vantajoso que o DCC ($n_f + n_a = 2^k + 2k = 14$), haja vista demandar apenas 12 combinações;
- b) Quando $k = 4$, o Box-Behnken [$2k(k - 1) = 24$] e o DCC ($n_f + n_a = 2^k + 2k = 24$) demandam o mesmo número de combinações;
- c) Quando $k = 5$, o DCC é mais recomendável que o Box-Behnken [$2k(k - 1) = 40$], haja vista que no DCC é possível utilizar, na parte fatorial, as combinações do fatorial fracionário 2^{5-1} , ou seja, DCC ($n_f + n_a = 2^{k-1} + 2k = 16 + 10 = 26$).

Uma grande desvantagem do delineamento Box-Behnken é quando existe a realização da análise rígida, isto é, a otimização não linear restrita, pois, na obtenção do ponto estacionário do modelo estimado $\hat{Y} = \hat{a}_o + x'_v \hat{a}_v + x'_v A x_v$, existe a restrição $x'_v x_v \leq \alpha^2$. No caso do DCC rotacionado, ocorre que $\alpha = \sqrt[4]{2^k}$, no qual o raio aumenta com o aumento do número de fatores. Já no delineamento Box-Behnken, é igual $\alpha = \sqrt{2}$, não sendo influenciado com o aumento do número de fatores, o que limita a análise de otimização não linear do planejamento Box-Behnken, haja vista o valor $\alpha = \sqrt{2}$ ser baixo, isto é, a sua região experimental é mais restrita devido à condição $-1 \leq x_i \leq +1$.

Exercício 11.4.1. Seja considerado um ensaio de adubação de uma espécie florestal, considerando os fatores N , P e K , com as dosagens 0, 60 e 120 kg por parcela. Na tabela 11.4.2, constam as produções de madeira em m^3 por parcela.

Tabela 11.4.2. Produção de madeira m^3 por parcela.

	x_1	x_2	x_3	X_1	X_2	X_3	Prod.
1	-1	-1	0	0	0	60	47,3
2	+1	-1	0	120	0	60	45,2
3	-1	+1	0	0	120	60	58,1
4	+1	+1	0	120	120	60	57,0
5	-1	0	-1	0	60	0	50,7
6	+1	0	-1	120	60	0	52,5

Tabela 11.4.2. Produção de madeira m³ por parcela (continuação).

	x_1	x_2	x_3	X_1	X_2	X_3	Prod.
7	-1	0	+1	0	60	120	57,3
8	+1	0	+1	120	60	120	52,2
9	0	-1	-1	60	0	0	45,8
10	0	+1	-1	60	120	0	52,1
11	0	-1	+1	60	0	120	42,2
12	0	+1	+1	60	120	120	58,8
13	0	0	0	60	60	60	55,1
14	0	0	0	60	60	60	56,3
15	0	0	0	60	60	60	57,1
16	0	0	0	60	60	60	61,8

Teste para verificar a existência de curvatura

$$SQ_{CURVATURA} = \frac{n_f \times n_c (\bar{y}_f - \bar{y}_c)^2}{n_f + n_c} = \frac{12 \times 4 \times (51,6 - 57,58)^2}{12 + 4}$$

$$SQ_{CURVATURA} = \frac{n_f \times n_c (\bar{y}_f - \bar{y}_c)^2}{n_f + n_c} = \frac{48 \times 35,7604}{16} = 107,2812$$

$\bar{y}_f = 51,6$ é o valor médio estimado dos valores dos pontos fatoriais;

$\bar{y}_c = 57,58$ é o valor médio estimado dos valores dos pontos centrais.

Então:

$$F = \frac{SQ_{CURVATURA}}{\hat{\sigma}^2} = \frac{107,2812}{8,61} = 12,46$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n_c} (y_{ci} - \bar{y}_c)^2}{n_c - 1} = 8,61 \quad F_{1;3;0,05} = 10,13$$

Em que $\hat{\sigma}^2 = 8,61$ é a estimativa do erro experimental, obtida a partir dos quatro valores dos pontos centrais. Comparando o valor calculado da estatística $F = 12,46$ com $F_{1;3;0,05} = 10,13$, conclui-se a existência da curvatura na região central.

A tabela 11.4.3 apresenta os resultados da análise considerando o modelo quadrático de segunda ordem. Verifica-se na tabela 11.4.4 que os efeitos lineares e quadráticos foram significantes.

$$Y_i = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^k a_{ii} x_i^2 + \sum_i \sum_j a_{ij} x_i x_j + e_i \quad i < j$$

Tabela 11.4.3. Estimativas dos parâmetros do modelo quadrático no SAS.

Parâmetros	G.L	Estimativas s/codificação	Erro padrão	Valor “t”	Pr > t	Estimativas c/codificação
Intercepto	1	43,662500	2,492990	17,51	< 0,0001	57,575000
N	1	0,048125	0,045169	1,07	0,3277	-0,812500
P	1	0,199792	0,045169	4,42	0,0045	5,687500
K	1	0,115000	0,045169	2,55	0,0437	1,175000
N×N	1	-0,000309	0,000295	-1,05	0,3356	-1,112500
P×N	1	0,000069444	0,000295	0,24	0,8219	0,250000
P×P	1	-0,001267	0,000295	-4,29	0,0051	-4,562500
K×N	1	-0,000479	0,000295	-1,62	0,1558	-1,725000
K×P	1	0,000715	0,000295	2,42	0,0517	2,575000
K×K	1	-0,000913	0,000295	-3,09	0,0213	-3,287500

Tabela 11.4.4. Resultado dos testes para os efeitos lineares, quadráticos e interações.

Regressão	G.L.	S.Q.	F	Pr > F
Linear	3	275,107500	20,29	0,0015
Quadrática	3	131,446875	9,69	0,0102
Interações	3	38,675000	2,85	0,1271
Total	9	445,229375	10,94	0,0044

Seja equação, considerando a seguinte codificação:

$$x_i = \frac{X_i - \bar{X}}{q} = \frac{X_i - 60}{60} \rightarrow X_i = 60x_i + 60$$

$$\hat{Y}_i = 57,575 - 0,8125x_1 + 5,6875x_2 + 1,175x_3 - 1,1125x_1^2 - 4,5625x_2^2 - 3,2875x_3^2 + 0,25x_1x_2 - 1,725x_1x_3 + 2,575x_2x_3$$

a) Matriz Hessiana A

$$A = \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12}/2 & \hat{a}_{13}/2 \\ \hat{a}_{12}/2 & \hat{a}_{22} & \hat{a}_{23}/2 \\ \hat{a}_{13}/2 & \hat{a}_{23}/2 & \hat{a}_{33} \end{bmatrix} = \begin{bmatrix} -1,1125 & 0,125 & -0,8625 \\ 0,125 & -4,5625 & 1,2875 \\ -0,8625 & 1,2875 & -3,2875 \end{bmatrix}$$

b) Obtenção do estacionário (x_{v0})

$$x_{v0} = -\frac{1}{2} A^{-1} \hat{a}_v = -\frac{1}{2} \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12}/2 & \hat{a}_{13}/2 \\ \hat{a}_{12}/2 & \hat{a}_{22} & \hat{a}_{23}/2 \\ \hat{a}_{13}/2 & \hat{a}_{23}/2 & \hat{a}_{33} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{bmatrix}$$

$$x_{v0} = -\frac{1}{2} \begin{bmatrix} -1,1125 & 0,125 & -0,8625 \\ 0,125 & -4,5625 & 1,2875 \\ -0,8625 & 1,2875 & -3,2875 \end{bmatrix}^{-1} \times \begin{bmatrix} -0,8125 \\ 5,6875 \\ 1,175 \end{bmatrix}$$

$$x_{v0} = \begin{bmatrix} -0,824577 \\ 0,8006597 \\ 0,7086073 \end{bmatrix} \rightarrow X_{v0} = \begin{bmatrix} 60 \times -0,824577 + 60 \\ 60 \times 0,8006597 + 60 \\ 60 \times 0,7086073 + 60 \end{bmatrix} = \begin{bmatrix} 10,5254 \\ 108,0396 \\ 102,5164 \end{bmatrix}$$

$X'_{v0} = [10,5254 \ 108,0396 \ 102,5164]$: vetor apresentando às doses de N, P e K, em kg por parcela, no ponto estacionário.

A produção no ponto estacionário é igual a $\hat{Y}_0 = 60,603167\text{m}^3$ por parcela. É importante observar que o ponto estacionário está dentro da região experimental, pois foram usadas as doses 0, 60, 120 kg por parcela.

c) Estudo da natureza do ponto estacionário (x_{v0})

$$\det(A - \lambda I) = \left\{ A = \begin{bmatrix} -1,1125 & 0,125 & -0,8625 \\ 0,125 & -4,5625 & 1,2875 \\ -0,8625 & 1,2875 & -3,2875 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} = 0$$

$$\hat{\lambda}_v = \begin{bmatrix} -0,782963 \\ -2,742429 \\ -5,437108 \end{bmatrix}$$

Dado que a matriz Hessiana A possui todos os autovalores negativos, conclui-se que A é definida negativa, logo $X'_{v0} = [10,5254 \ 108,0396 \ 102,5164]$ é um ponto de máximo. O ponto de máximo está dentro do espaço amostral 0, 60, 120 kg por parcela e apresentando a produção máxima de $\hat{Y}_0 = 60,603167 \text{ m}^3$ por parcela.

Tabela 11.4.5. Valores críticos, autovalores e autovetores no SAS.

Fator	Valores Críticos		Autovalor	Autovetores		
	c/codific.	s/codific.		N	P	K
N	-0,824577	10,525410	-0,782963	0,925317	-0,094495	-0,36723
P	0,800660	108,039582	-2,742429	0,355043	0,556009	0,75153
K	0,708607	102,516440	-5,437108	-0,133169	0,825787	-0,54803
Valor estimado do ponto estacionário: 60,603167			O ponto estacionário é de máximo			

d) A relação entre as variáveis $x'_v = [x_1, x_2, x_3]$ e $u'_v = [u_1, u_2, u_3]$:

Dada a matriz M dos autovetores normalizados correspondentes aos autovalores da matriz A :

$$\lambda'_v = [-0,782963 \ -2,742429 \ -5,437108],$$

$$M = \begin{bmatrix} 0,925317 & 0,3550426 & -0,133169 \\ -0,094495 & 0,5560092 & 0,8257872 \\ -0,367233 & 0,7515308 & -0,548034 \end{bmatrix}$$

Então:

$$u_v = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = M' z_v = \begin{bmatrix} 0,925317 & -0,094495 & -0,367233 \\ 0,3550426 & 0,5560092 & 0,7515308 \\ -0,133169 & 0,8257872 & -0,548034 \end{bmatrix} \times \begin{bmatrix} x_1 + 0,824577 \\ x_2 - 0,800660 \\ x_3 - 0,708607 \end{bmatrix}$$

Conclui-se que para N, P, K , a dosagem que maximiza a produção, sem considerar a parte econômica, é (10,5254 102,5164 108,0396), kg por parcela, a qual está dentro do intervalo estudado 0, 60, 120 kg por parcela, correspondendo à produção máxima de madeira igual a 60,6032 m^3 por parcela. Programa SAS, **Prog38**, anexo no apêndice B.

Como visto em aplicações anteriores, pode-se obter a produção máxima, considerando a dose economicamente aconselhável.

11.5. EXEMPLO DE APLICAÇÃO DE UM FATORIAL $(\frac{1}{5})5^3$

Exercício 11.5.1. Seja considerado um experimento de adubação para uma espécie florestal (dados simulados), em um fatorial fracionário $(\frac{1}{5})5^3$ com N, P e K com as doses 0, 30, 60, 90 e 120 kg ha^{-1} . A variável resposta observada foi o volume de madeira em $\text{m}^3 \text{ ha}^{-1}$ no final de rotação, usando o planejamento apresentado por Conagin e Jorge (1977), o qual é constituído por apenas 25 das 125 combinações do fatorial completo, em que os seguintes tratamentos constituem um único bloco.

111	134	152	125	143
254	222	245	213	231
342	315	333	351	324
435	453	421	444	412
523	541	514	532	555

Tabela 11.5.1. Volume de madeira em m³ ha⁻¹.

N	P	K	Prod.	N	P	K	Prod.	N	P	K	Prod.
0	0	0	60,5	0	60	90	148,4	0	30	120	108,7
30	30	30	134,5	60	30	90	158,8	30	60	0	155,7
60	60	60	168,9	90	60	120	184,1	60	90	30	186,7
90	90	90	199,9	120	90	0	144,5	120	60	30	184,6
120	120	120	183,0	0	120	30	122,7	0	90	60	155,7
30	90	120	158,7	30	0	60	108,1	30	120	90	179,1
60	120	0	142,1	90	120	60	165,3	60	0	120	135,9
90	0	30	126,7	120	0	90	147,1	90	30	0	115,3
120	30	60	140,4								

$$Y_i = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^k a_{ii} x_i^2 + \sum_i \sum_j a_{ij} x_i x_j + e_i \quad i < j$$

a) Estimativas dos parâmetros do modelo no SAS

Tabela 11.5.2. Estimativas dos parâmetros do modelo quadrático (SAS).

Parâmetros	G.L.	Estimativas s/codificação	Erro padrão	Valor “t”	Pr > t	Estimativas c/codificação
Intercepto	1	63,189060	11,190946	5,65	<0,0001	179,838793
N	1	0,822645	0,272988	3,01	0,0087	18,496000
P	1	1,480270	0,277492	5,33	<0,0001	24,624000
K	1	0,611609	0,277492	2,20	0,0436	15,308000
N×N	1	-0,005046	0,001800	-2,80	0,0134	-18,165176
P×N	1	-0,001097	0,001690	-0,65	0,5260	-3,950263
P×P	1	-0,008194	0,001925	-4,26	0,0007	-29,499094
K×N	1	0,002616	0,001690	1,55	0,1424	9,417885
K×P	1	-0,000345	0,001668	-0,21	0,8387	-1,243769
K×K	1	-0,004106	0,001925	-2,13	0,0499	-14,781316

A tabela 11.5.2 apresenta a análise para o modelo quadrático de segunda ordem.

$$\hat{Y}_i = 179,838793 + 18,496x_1 + 24,624x_2 + 15,308x_3 - 18,165176x_1^2 - 29,499094x_2^2 - 14,781316x_3^2 - 3,950263x_1x_2 + 9,417885x_1x_3 - 1,243769x_2x_3$$

$$x_i = \frac{X_i - \bar{X}}{q} = \frac{X_i - 60}{60} \rightarrow X_i = 60x_i + 60$$

b) Testes para efeitos lineares, quadráticos e interações

Ajustando o modelo quadrático, verifica-se, na tabela 11.5.3, que os efeitos lineares e quadráticos foram significantes.

Tabela 11.5.3. Resultado dos testes para os efeitos lineares, quadráticos e interações.

Regressão	G.L.	S.Q.	F	Pr > F
Linear	3	14785,0	27,31	< 0,0001
Quadrático	3	5495,960714	10,15	0,0007
Interações	3	505,593459	0,93	0,4486
Total	9	20786,0	12,80	< 0,0001

c) Matriz Hessiana A

$$A = \begin{bmatrix} -18,165176 & -1,975132 & 4,708943 \\ -1,975132 & -29,499094 & -0,621885 \\ 4,708943 & -0,621885 & -14,781316 \end{bmatrix}$$

d) Obtenção do estacionário (x_{v0})

$$x_{v0} = -\frac{1}{2}A^{-1}\hat{a}_v = -\frac{1}{2} \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12}/2 & \hat{a}_{13}/2 \\ \hat{a}_{12}/2 & \hat{a}_{22} & \hat{a}_{23}/2 \\ \hat{a}_{13}/2 & \hat{a}_{23}/2 & \hat{a}_{33} \end{bmatrix}^{-1} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{bmatrix}$$

Logo:

$$x_{v0} = -\frac{1}{2} \begin{bmatrix} -18,165176 & -1,975132 & 4,708943 \\ -1,975132 & -29,499094 & -0,621885 \\ 4,708943 & -0,621885 & -14,781316 \end{bmatrix}^{-1} \times \begin{bmatrix} 18,496 \\ 24,624 \\ 15,308 \end{bmatrix}$$

$$x_{v0} = \begin{bmatrix} 0,654493 \\ 0,358553 \\ 0,711235 \end{bmatrix} \rightarrow X_{v0} = \begin{bmatrix} 60 \times 0,654493 + 60 \\ 60 \times 0,358553 + 60 \\ 60 \times 0,711235 + 60 \end{bmatrix} = \begin{bmatrix} 99,2696 \\ 81,5132 \\ 102,6741 \end{bmatrix}$$

$X'_{v0} = [99,2696 \ 81,5132 \ 102,6741]$ é o vetor com as doses de N , P e K , em kg ha^{-1} , no ponto estacionário. Nesse ponto, a produção é igual a $\hat{Y}_0 = 195,749838 \text{ m}^3 \text{ ha}^{-1}$. É importante observar que o ponto estacionário está dentro da região experimental, pois foram usadas as doses 0, 30, 60, 90, 120 kg ha^{-1} .

e) Estudo da natureza do ponto estacionário (x_{v0})

Tabela 11.5.4. Valores críticos, autovalores e autovetores no SAS.

Factor	Valores Críticos		Autovalores	Autovetores		
	c/codific.	s/codific.		N	P	K
N	0,654493	99,269565	-11,319452	0,582055	-0,090879	0,808055
P	0,358553	81,513172	-21,290437	0,794752	-0,146610	-0,588961
K	0,711235	102,674105	-29,835697	0,171994	0,985011	-0,013109
Valor estimado no ponto estacionário: 195,749838			O ponto estacionário é de máximo			

$$\det(A - \lambda I) = \left\{ A = \begin{bmatrix} -18,165176 & -1,975132 & 4,708943 \\ -1,975132 & -29,499094 & -0,621885 \\ 4,708943 & -0,621885 & -14,781316 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} = 0$$

$$\hat{\lambda}_v = \begin{bmatrix} -11,319452 \\ -21,290437 \\ -29,835697 \end{bmatrix}$$

Dado que a matriz Hessiana A possui todos os autovalores negativos, então A é definida negativa, logo $X'_{v0} = [99,2696 \ 81,5132 \ 102,6741]$ é um ponto de máximo. O ponto de máximo está dentro do espaço amostral (0, 30, 60, 90, 120), kg ha^{-1} , apresentando a produção máxima de $\hat{Y}_0 = 195,749838 \text{ m}^3 \text{ ha}^{-1}$.

f) A relação entre as variáveis $x'_v = [x_1, x_2, x_3]$ e $u'_v = [u_1, u_2, u_3]$

$\lambda'_v = [-11,319452 \ -21,290437 \ -29,835697]$ é o vetor dos autovalores da matriz A .

Seja considerada a matriz M dos autovetores normalizados correspondentes aos autovalores da matriz A :

$$M = \begin{bmatrix} 0,582055 & 0,794752 & 0,171994 \\ -0,090879 & -0,146610 & 0,985011 \\ 0,808055 & -0,588961 & -0,013109 \end{bmatrix}$$

Então:

$$u_v = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = M z_v = \begin{bmatrix} 0,582055 & -0,090879 & 0,808055 \\ 0,794752 & -0,146610 & -0,588961 \\ 0,171994 & 0,985011 & -0,013109 \end{bmatrix} \times \begin{bmatrix} x_1 - 0,654493 \\ x_2 - 0,358553 \\ x_3 - 0,711235 \end{bmatrix}$$

Conclui-se que, para N, P, K , a dosagem que maximiza a produção, sem considerar a parte econômica é (99,2696 102,6741 81,5132) kg ha⁻¹, a qual está dentro do intervalo estudado (0, 30, 60, 90, 120) kg ha⁻¹, correspondendo à produção máxima de madeira igual a 195,749838 m³ ha⁻¹. A análise foi realizada pelo programa **Prog39** encontrado no apêndice *B* em anexo.

Como visto em aplicações anteriores, nesse exemplo, pode-se obter a produção máxima considerando a dose economicamente aconselhável.

Pimentel Gomes e Garcia (2002) efetuaram a análise das estimativas de produção do ensaio de adubação de milho, no estado de São Paulo, de um fatorial fracionário $(\frac{1}{5})5^3$ com N, P e K , apresentado por Conagin e Jorge (1982), em que as doses de cada nutriente foi 0, 25, 50, 75 e 100 kg ha⁻¹. A análise apresentou como resultado a seguinte equação:

$$\hat{Y}_i = 1.995,42 + 13,4423N + 32,1975P + 6,3554K - 0,074686N^2 - 0,206491P^2 - 0,057780K^2 - 0,019847NP + 0,065812NK + 0,023599PK$$

Por outro lado, essa equação apresentou a produção máxima de $\hat{Y}_{Max} = 4.767,6$ kg ha⁻¹ no ponto ($N = 148$ kg, $P = 80$ kg, $K = 156$ kg), por hectare de nutriente. Entretanto, como o ponto está fora do cubo explorado, em que o intervalo utilizado foi (0, 25, 50, 75, 100), referidos autores obtiveram um máximo absoluto na fronteira desse cubo, ou seja, nas suas faces. O procedimento utilizado foi substituir os valores $N = K = 100$ na equação $\hat{Y}$, resultando em:

$$\hat{Y}^* = 3.308,65 + 32,5727P - 0,206491P^2$$

Derivando em relação a P e igualando a zero, resultou a dose $P^* = 79$ kg ha⁻¹. Assim, o ponto de máximo absoluto, de acordo com Pimentel Gomes e Garcia (2002), deve dar-se no ponto:

($N^* = 100$ kg, $P^* = 79$ kg, $K^* = 100$ kg), em que $\hat{Y}_{Max}^* = 4.593,2$ kg ha⁻¹.

11.6. BIBLIOGRAFIA

BOX, G.E.P.; BEHNKEN, D.W. Some New Three Level Designs for the Study of Quantitative Variables. **Technometrics**, 2, 455-475, 1960.

CARPINETTI, L. C. R. **Planejamento e análise de experimentos**. Escola de Engenharia de São Carlos. Departamento de Engenharia de Produção. USP, 2009. 222 p.

CONAGIN, A.; JORGE, J. P. N. Delineamentos $(\frac{1}{5})5^3$. **Bragantia**, 36: 23-58. 1977.

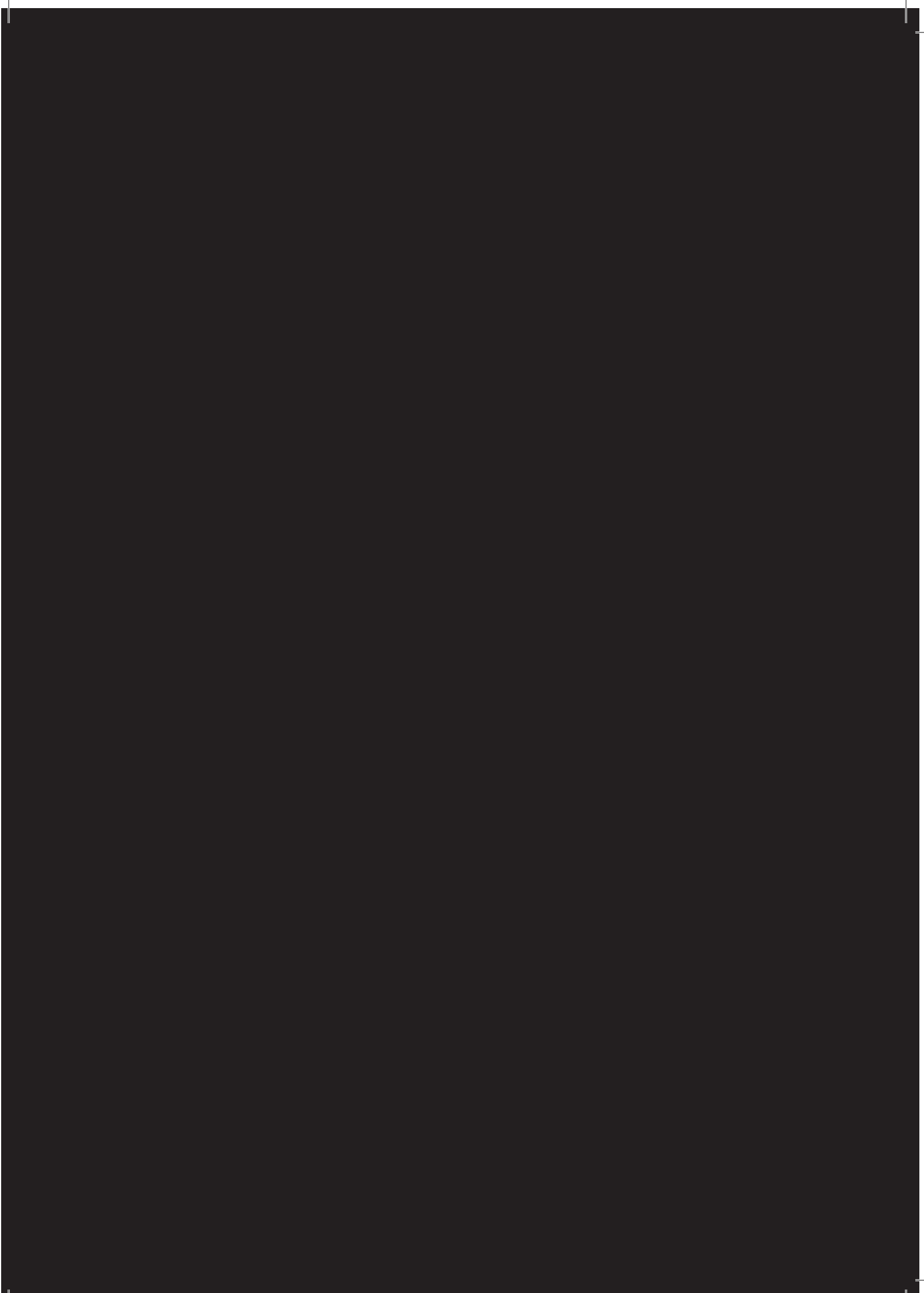
CONAGIN, A.; JORGE, J. P. N. Delineamentos $(\frac{1}{5})5^3$ em blocos. **Bragantia**, 41: 155-168, 1982.

LIMA, C. N. **Escolha de tratamentos e blocagem otimizados na construção de fatoriais fracionários**. Tese de doutorado. Programa de pós-graduação em estatística e experimentação agropecuária. UFLA. Lavras – MG, 2009. 147 p.

MENDONÇA, L. A. **Desempenho do delineamento composto central em experimento com alto coeficiente de variação**. 80 p. Dissertação de mestrado. Programa de pós-graduação em estatística aplicada e biometria. UFV. Viçosa, 2012.

PIMENTEL GOMES, F.; GARCIA, C. H. **Estatística aplicada a experimentos agrônômicos e florestais**. Piracicaba: FEALQ. (Biblioteca de Ciências Agrárias Luiz de Queiroz - Volume 11), 2002. 309 p.

SILVA, M. A. **Critérios compostos para delineamentos ótimos robustos**. Dissertação de mestrado. Pós-graduação em Estatística e Experimentação Agronômica. Universidade Estadual Paulista “Júlio de Mesquita”. Botucatu-SP, 2014. 162 p.



APÊNDICE A

DISTRIBUIÇÕES ESTATÍSTICAS

Tabela A1. Áreas da distribuição normal padrão.

Tabela A2. Distribuição “ t ” de Student.

Tabela A3. Distribuição F de Snedecor para $\alpha = 0,05$.

Tabela A4. Distribuição F de Snedecor para $\alpha = 0,01$.

Tabela A5. Distribuição de χ^2 .

Tabela A6. Teste de Kolmogorov-Smirnov: valores críticos de D para uma amostra.

Tabela A7 – Teste de Cochran: homogeneidade de variâncias ($\alpha = 0,05$).

Tabela A8. Teste de Dunnett: valores de (t_d) para $\alpha = 0,05$.

Tabela A9. Tabela de F máximo ou H de Hartley para $\alpha = 0,05$.

Tabela A10. Tabela de F máximo ou H de Hartley para $\alpha = 0,01$.

Tabela A11. Limites superiores para o teste de Lilliefors.

Tabela A12. Tabela de “ t ” (bilaterais) para o teste de Bonferroni para $\alpha = 0,05$.

Tabela A13. Valores de “ t ” (bilaterais) para o teste de Bonferroni para $\alpha = 0,01$.

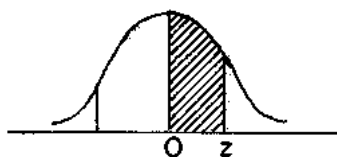
Tabela A14. Teste de Tukey: valores da amplitude total estudentizada para $\alpha = 0,05$.

Tabela A15. Teste de Tukey: valores da amplitude total estudentizada para $\alpha = 0,01$.

Tabela A16. Teste de Duncan: valores da amplitude total estudentizada para $\alpha = 0,05$.

Tabela A17. Teste de Duncan: valores da amplitude total estudentizada para $\alpha = 0,01$.

Tabela A18. Quantis de ordem ($p = 1 - \alpha$) para o teste de Shapiro-Wilk (W_p)

**Tabela A1.** Áreas da distribuição normal padrão.

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

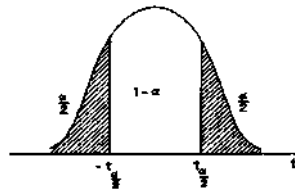
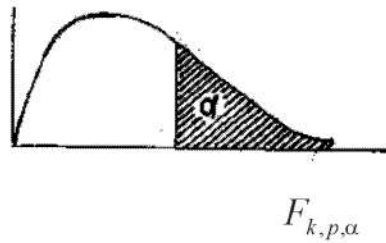
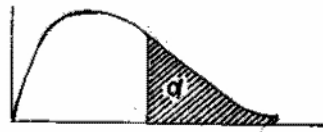


Tabela A2. Distribuição “ t ” de Student ($k = \text{ngl}$).

$k \backslash \alpha$	0,50	0,25	0,10	0,05	0,025	0,01	0,005
1	1,00000	2,4142	6,3138	12,706	25,542	63,657	127,32
2	0,81650	1,6036	2,9200	4,3127	6,2053	9,9248	14,089
3	0,76489	1,4226	2,3534	3,1825	4,1765	5,8409	7,4533
4	0,74070	1,3444	2,1318	2,7764	3,4954	4,6041	5,5976
5	0,72669	1,3009	2,0150	2,5706	3,1634	4,0321	4,7733
6	0,71756	1,2733	1,9432	2,4469	2,9687	3,7074	4,3168
7	0,71114	1,2543	1,8946	2,3646	2,8412	3,4995	4,0293
8	0,70639	1,2403	1,8595	2,3060	2,7515	3,3554	3,8325
9	0,70272	1,2297	1,8331	2,2622	2,6850	3,2498	3,6897
10	0,69981	1,2213	1,8125	2,2281	2,6338	3,1693	3,5814
11	0,69745	1,2145	1,7959	2,2010	2,5931	3,1058	3,4966
12	0,69548	1,2089	1,7823	2,1788	2,5600	3,9545	3,4284
13	0,69384	1,2041	1,7709	2,1604	2,5326	3,0123	3,3725
14	0,69200	1,2001	1,7613	2,1448	2,5096	5,9768	3,3257
15	0,69120	1,1967	1,7530	2,1315	2,4899	2,9467	3,2860
16	0,69013	1,1937	1,7459	2,1199	2,4729	2,9208	3,2520
17	0,68919	1,1910	1,7396	2,1098	2,4581	2,8982	3,2225
18	0,68837	1,1887	1,7341	2,1009	2,4450	2,8784	3,1966
19	0,68763	1,1866	1,7291	2,0930	2,4334	2,8609	3,1737
20	0,68696	1,1848	1,7247	2,0860	2,4231	2,8453	3,1534
21	0,68635	1,1831	1,7207	2,0796	2,4138	2,8314	3,1352
22	0,68580	1,1816	1,7171	2,0739	2,4055	2,8188	3,1188
23	0,68531	1,1802	1,7139	2,0687	2,3979	2,8073	3,1040
24	0,68485	1,1789	1,7109	2,0639	2,3910	2,7969	3,0905
25	0,68443	1,1777	1,7081	2,0595	2,3846	2,7874	3,0782
26	0,68405	1,1766	1,7056	2,0555	2,3788	2,7787	3,0669
27	0,68370	1,1757	1,7033	2,0518	2,3734	2,7707	3,0565
28	0,68335	1,1748	1,7011	2,0484	2,3685	2,7633	3,0469
29	0,68304	1,1739	1,6991	2,0452	2,3638	2,7564	3,0380
30	0,68276	1,1731	1,6973	2,0423	2,3596	2,7500	3,0298
40	0,68066	1,1673	1,6839	2,0211	2,3289	2,7045	2,9712
60	0,67862	1,1616	1,6707	2,0003	2,2991	2,6603	2,9146
120	0,67656	1,1559	1,6577	1,9799	2,2699	2,6174	2,8599
∞	0,67449	1,1503	1,6449	1,9600	2,2414	2,5758	2,8070

Tabela A3. Distribuição F de Snedecor para $\alpha = 5\%$.

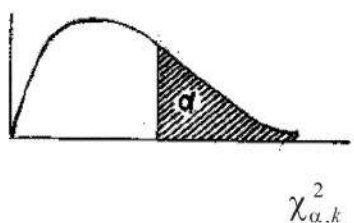
$k \backslash p$	1	2	3	4	5	6	7	8	9	10	20	30	120
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	248,0	250,1	253,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,45	19,46	19,49
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,66	8,62	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,80	5,75	5,66
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,56	4,50	4,40
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,87	3,81	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,44	3,38	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,15	3,08	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	2,94	2,86	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,77	2,70	2,58
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,65	2,57	2,45
12	4,75	3,89	3,49	3,26	3,11	3,00	2,90	2,85	2,80	2,75	2,54	2,47	2,34
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,46	2,38	2,25
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,39	2,31	2,18
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,33	2,25	2,11
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,28	2,19	2,06
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,23	2,15	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,19	2,11	1,97
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,16	2,07	1,93
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,12	2,04	1,90
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,10	2,01	1,87
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,07	1,98	1,84
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,05	1,96	1,81
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,03	1,94	1,79
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	1,93	1,84	1,68
40	4,08	3,23	2,84	2,61	2,45	2,34	2,21	2,18	2,12	2,08	1,84	1,74	1,58
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,75	1,65	1,47
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,66	1,55	1,35
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,57	1,46	1,22



$$F_{k,p,\alpha}$$

Tabela A4. Distribuição F de Snedecor para $\alpha = 1\%$.

$k \backslash p$	1	2	3	4	5	6	7	8	9	10	20	30	120
1	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6209	6261	6339
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,45	99,47	99,49
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	26,69	26,50	26,22
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,02	13,84	13,56
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,55	9,38	9,11
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,40	7,23	6,97
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,16	5,99	5,74
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,36	5,20	4,95
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	4,81	4,65	4,40
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,41	4,25	4,00
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,10	3,94	3,69
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	3,86	3,70	3,45
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,66	3,51	3,25
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,51	3,35	3,09
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,37	3,21	2,96
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,26	3,10	2,84
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,16	3,00	2,75
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,08	2,92	2,66
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,00	2,84	2,58
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	2,94	2,78	2,52
21	8,02	5,78	4,82	4,37	4,04	3,81	3,64	3,51	3,40	3,31	2,88	2,72	2,46
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	2,83	2,67	2,40
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	2,78	2,62	2,35
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	2,74	2,58	2,31
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,70	2,54	2,27
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,66	2,50	2,23
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,63	2,47	2,20
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,60	2,44	2,17
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,57	2,41	2,14
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,55	2,39	2,11
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,37	2,20	1,92
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,20	2,03	1,73
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,03	1,86	1,53
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	1,88	1,70	1,32

Tabela A5. Distribuição de χ^2 .

$\alpha \backslash k$	0,995	0,990	0,975	0,95	0,90	0,75	0,50	0,50	0,10	0,05	0,025	0,01	0,005
1	0,000	0,0002	0,0010	0,0039	0,0158	0,102	0,455	1,32	2,71	3,84	5,02	6,63	7,88
2	0,010	0,0001	0,0506	0,103	0,211	0,575	1,39	2,77	4,61	5,99	7,38	9,21	10,6
3	0,0717	0,115	0,216	0,352	0,584	1,021	2,37	4,11	6,25	7,81	9,25	11,3	12,8
4	0,207	0,297	0,484	0,711	1,06	1,92	3,36	5,39	7,78	9,49	11,1	13,3	14,9
5	0,412	0,554	0,831	1,15	1,61	2,67	4,35	6,63	9,24	11,1	12,8	15,1	16,7
6	0,676	0,872	1,24	1,64	2,20	3,45	5,35	7,84	10,6	12,6	14,4	16,8	18,5
7	0,989	1,24	1,69	2,17	2,83	4,25	6,35	9,04	12,0	14,1	16,0	18,5	20,3
8	1,34	1,65	2,18	2,73	3,49	5,07	7,34	10,2	13,4	15,5	17,5	20,1	22,0
9	1,73	2,09	2,70	3,33	4,17	5,90	8,34	11,4	14,7	16,9	19,0	21,7	23,6
10	2,16	2,56	3,25	3,94	4,87	6,74	9,34	12,5	16,0	18,3	20,5	23,2	25,2
11	2,60	3,05	3,82	4,57	5,58	7,58	10,3	13,7	17,3	19,7	21,9	24,7	26,8
12	3,07	3,57	4,40	5,23	6,30	8,44	11,3	14,8	18,5	21,0	23,3	26,2	28,3
13	3,57	4,11	5,01	5,89	7,04	9,30	12,3	16,0	19,8	22,4	24,7	27,7	29,8
14	4,07	4,66	5,63	6,57	7,79	10,2	13,3	17,1	21,1	23,7	26,1	29,1	31,3
15	4,60	5,23	6,23	7,26	8,55	11,0	14,3	18,2	22,3	25,0	27,5	30,6	32,8
16	5,14	5,80	6,91	7,96	9,31	11,9	15,3	19,4	23,5	26,3	28,4	32,0	34,3
17	5,70	6,41	7,56	8,67	10,1	12,8	16,3	20,5	24,8	27,6	30,2	33,4	35,7
18	6,26	7,01	8,23	9,39	10,9	13,7	17,3	21,6	26,0	28,9	31,5	34,8	37,2
19	6,84	7,63	8,91	10,1	11,7	14,6	18,3	22,7	27,2	30,1	32,9	36,2	38,6
20	7,43	8,26	9,59	10,9	12,4	15,5	19,3	23,8	28,4	31,4	34,2	37,6	40,0
21	8,03	8,90	10,3	11,6	13,2	16,3	20,3	24,9	29,6	32,7	35,5	38,9	41,4
22	8,64	9,54	11,0	12,3	14,0	17,2	21,3	26,0	30,8	33,9	36,8	40,5	42,8
23	9,26	10,2	11,7	13,1	14,8	18,1	22,3	27,1	32,0	35,2	38,1	41,6	44,2
24	9,89	10,9	12,4	13,8	15,7	19,0	23,3	28,2	33,1	36,4	39,4	43,0	45,6
25	10,5	11,5	13,1	14,6	16,5	19,9	24,3	29,3	34,4	37,7	40,6	44,3	46,9
26	11,2	12,2	13,8	15,4	17,3	20,8	25,3	30,4	35,6	38,9	41,9	45,6	48,3
27	11,8	12,9	14,6	16,2	18,1	21,7	26,3	31,5	36,7	40,1	43,2	47,0	49,6
28	12,5	13,6	15,3	16,9	18,9	22,7	27,3	32,6	37,9	41,3	44,5	48,3	51,0
29	13,1	14,3	16,0	17,7	19,8	23,6	28,3	33,7	39,1	42,6	45,7	49,6	52,5
30	13,8	15,0	16,8	18,5	20,6	24,5	29,3	34,8	40,3	43,8	47,0	50,9	53,7

Tabela A6. Teste de Kolmogorov-Smirnov: valores críticos de D para uma amostra.

α (nível de significância)					α (nível de significância)				
n	0,20	0,10	0,05	0,01	n	0,20	0,10	0,05	0,01
1	0,900	0,950	0,975	0,995	36	0,174	0,199	0,221	0,265
2	0,684	0,776	0,842	0,929	37	0,172	0,196	0,218	0,262
3	0,565	0,636	0,708	0,829	38	0,170	0,194	0,215	0,258
4	0,493	0,565	0,624	0,734	39	0,168	0,191	0,213	0,255
5	0,447	0,509	0,563	0,668	40	0,165	0,189	0,210	0,252
6	0,410	0,468	0,519	0,617	41	0,163	0,187	0,208	0,249
7	0,381	0,436	0,483	0,576	42	0,162	0,185	0,206	0,246
8	0,358	0,410	0,454	0,542	43	0,160	0,183	0,203	0,243
9	0,339	0,387	0,430	0,513	44	0,158	0,180	0,200	0,241
10	0,323	0,369	0,409	0,489	45	0,156	0,178	0,198	0,238
11	0,306	0,352	0,391	0,468	46	0,154	0,177	0,196	0,235
12	0,296	0,338	0,375	0,449	47	0,153	0,175	0,194	0,233
13	0,285	0,325	0,361	0,432	48	0,151	0,173	0,192	0,231
14	0,275	0,314	0,349	0,418	49	0,150	0,171	0,190	0,228
15	0,266	0,304	0,338	0,404	50	0,148	0,170	0,188	0,226
16	0,258	0,295	0,327	0,392	51	0,147	0,168	0,186	0,224
17	0,250	0,286	0,318	0,381	52	0,146	0,166	0,185	0,222
18	0,244	0,278	0,309	0,371	53	0,144	0,165	0,183	0,220
19	0,237	0,271	0,301	0,361	54	0,143	0,163	0,181	0,218
20	0,232	0,265	0,294	0,352	55	0,142	0,162	0,180	0,216
21	0,226	0,258	0,287	0,344	56	0,140	0,160	0,178	0,214
22	0,221	0,253	0,281	0,337	57	0,139	0,159	0,177	0,212
23	0,216	0,247	0,275	0,330	58	0,138	0,158	0,175	0,210
24	0,212	0,242	0,269	0,323	59	0,137	0,156	0,174	0,208
25	0,208	0,238	0,264	0,317	60	0,136	0,155	0,172	0,207
26	0,204	0,233	0,259	0,311	61	0,135	0,154	0,171	0,205
27	0,200	0,229	0,254	0,305	62	0,134	0,153	0,170	0,203
28	0,197	0,225	0,250	0,300	63	0,133	0,151	0,168	0,202
29	0,193	0,221	0,246	0,295	64	0,132	0,150	0,167	0,200
30	0,190	0,218	0,242	0,290	65	0,131	0,149	0,166	0,199
31	0,187	0,214	0,238	0,285	66	0,130	0,148	0,164	0,197
32	0,184	0,211	0,234	0,281	67	0,129	0,147	0,163	0,196
33	0,182	0,208	0,231	0,277	68	0,128	0,146	0,162	0,194
34	0,179	0,205	0,227	0,273	69	0,127	0,145	0,161	0,193
35	0,176	0,202	0,224	0,269	70	0,126	0,144	0,160	0,192
						$\frac{1,06}{\sqrt{n}}$	$\frac{1,21}{\sqrt{n}}$	$\frac{1,35}{\sqrt{n}}$	$\frac{1,62}{\sqrt{n}}$

Tabela A7. Teste de Cochran: homogeneidade de variâncias ($\alpha = 0,05$).

NT	J: Número de repetições em cada tratamento (NT: Número de tratamentos)									
	2	3	4	5	6	7	8	9	10	11
2	0,9985	0,9750	0,9392	0,9057	0,8772	0,8534	0,8332	0,8159	0,8010	0,7880
3	0,9669	0,8709	0,7977	0,7457	0,7071	0,6771	0,6530	0,6333	0,6167	0,6025
3	0,9065	0,7679	0,6841	0,6287	0,5895	0,5598	0,5365	0,5175	0,5017	0,4884
5	0,8413	0,6838	0,5931	0,5441	0,5065	0,4783	0,4564	0,4387	0,4241	0,4118
6	0,7808	0,6161	0,5321	0,4803	0,4447	0,4194	0,3980	0,3817	0,3682	0,3568
7	0,7271	0,5612	0,4800	0,4307	0,3974	0,3726	0,3535	0,3384	0,3299	0,3154
8	0,6798	0,5157	0,4377	0,3910	0,3595	0,3362	0,3185	0,3043	0,2926	0,2829
9	0,6385	0,4775	0,4027	0,3584	0,3286	0,3067	0,2901	0,2768	0,2659	0,2568
10	0,6020	0,4455	0,3733	0,3311	0,3029	0,2823	0,0666	0,2541	0,2439	0,2353
12	0,5410	0,3924	0,3264	0,2880	0,2624	0,2439	0,2299	0,2187	0,2098	0,2020
15	0,4709	0,3346	0,2758	0,2419	0,2195	0,2034	0,1911	0,1815	0,1736	0,1671
20	0,3894	0,2705	0,2205	0,1921	0,1735	0,1602	0,1501	0,1422	0,1357	0,1303
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374	0,1246	0,1216	0,1160	0,1113
30	0,2929	0,1980	0,1593	0,1377	0,1237	0,1137	0,1061	0,1002	0,0958	0,0921
40	0,2370	0,1576	0,1259	0,1082	0,0968	0,0887	0,0827	0,0780	0,0745	0,0713
60	0,1737	0,1131	0,0895	0,0765	0,0682	0,0625	0,0583	0,0552	0,0520	0,0497
120	0,0998	0,0632	0,0495	0,0419	0,0371	0,0337	0,0312	0,0292	0,0279	0,0266

Tabela A8. Teste de Dunnett, valores de (t_d), para o nível de 5 % (n_{GL} é o número de graus de liberdade do resíduo).

n_{GL}	Número de graus de liberdade de tratamentos ($I-1$)													
	1	2	3	4	5	6	7	8	9	10	11	12	15	20
5	2,57	3,03	3,29	3,48	3,62	3,73	3,82	3,90	3,97	4,03	4,09	4,14	4,26	4,42
6	2,45	2,86	3,10	3,26	3,39	3,49	3,57	3,64	3,71	3,76	3,81	3,86	3,97	4,11
7	2,36	2,75	2,97	3,12	3,24	3,33	3,41	3,47	3,53	3,58	3,63	3,67	3,78	3,91
8	2,31	2,67	2,88	3,02	3,13	3,22	3,29	3,35	3,41	3,46	3,50	3,54	3,64	3,76
9	2,26	2,61	2,81	2,95	3,05	3,14	3,20	3,26	3,32	3,36	3,40	3,44	3,53	3,65
10	2,23	2,57	2,76	2,89	2,99	3,07	3,14	3,19	3,24	3,29	3,33	3,36	3,45	3,57
11	2,20	2,53	2,72	2,84	2,94	3,02	3,08	3,14	3,19	3,23	3,27	3,30	3,39	3,50
12	2,18	2,50	2,68	2,81	2,90	2,98	3,04	3,09	3,14	3,18	3,22	3,25	3,34	3,45
13	2,16	2,48	2,65	2,78	2,87	2,94	3,00	3,06	3,10	3,14	3,18	3,21	3,29	3,40
14	2,14	2,46	2,63	2,75	2,84	2,91	2,97	3,02	3,07	3,11	3,14	3,18	3,26	3,36
15	2,13	2,44	2,61	2,73	2,82	2,89	2,95	3,00	3,04	3,08	3,12	3,15	3,23	3,33
16	2,12	2,42	2,59	2,71	2,80	2,87	2,92	2,97	3,02	3,06	3,09	3,12	3,20	3,30
17	2,11	2,41	2,58	2,69	2,78	2,85	2,90	2,95	3,00	3,03	3,07	3,10	3,18	3,27
18	2,10	2,40	2,56	2,68	2,76	2,83	2,89	2,94	2,98	3,01	3,05	3,08	3,16	3,25
19	2,09	2,39	2,55	2,66	2,75	2,81	2,87	2,92	2,96	3,00	3,03	3,06	3,14	3,23
20	2,09	2,38	2,54	2,65	2,73	2,80	2,86	2,90	2,95	2,98	3,02	3,05	3,12	3,22
24	2,06	2,35	2,51	2,61	2,70	2,76	2,81	2,86	2,90	2,94	2,97	3,00	3,07	3,16
30	2,04	2,32	2,47	2,58	2,66	2,72	2,77	2,81	2,86	2,89	2,92	2,95	3,02	3,11
40	2,02	2,29	2,44	2,54	2,62	2,68	2,73	2,77	2,81	2,85	2,87	2,90	2,97	3,06
60	2,00	2,27	2,41	2,51	2,58	2,64	2,69	2,73	2,77	2,80	2,83	2,86	2,92	3,00
120	1,98	2,24	2,38	2,47	2,55	2,60	2,65	2,69	2,73	2,76	2,79	2,81	2,87	2,95
∞	1,96	2,21	2,35	2,44	2,51	2,57	2,61	2,65	2,69	2,72	2,74	2,77	2,83	2,91

Tabela A9. Tabela de F máximo ou H de Hartley para $\alpha = 0,05$.

	Número de estimativas de variâncias ($\alpha = 0,05$). GL = $J-1$										
GL	2	3	4	5	6	7	8	9	10	11	12
3	15,40	27,80	39,20	50,70	62,00	72,90	83,50	93,90	104,0	114,0	124,0
4	9,60	15,50	20,60	25,20	29,50	33,60	37,50	41,10	44,60	48,00	51,40
5	7,15	10,80	13,70	16,30	18,70	20,80	22,90	24,70	26,50	28,20	29,90
6	5,82	8,38	10,40	12,10	13,70	15,00	16,30	17,50	18,60	19,70	20,70
7	4,99	6,94	8,44	9,70	10,80	11,80	12,70	13,50	14,30	15,10	15,80
8	4,43	6,00	7,18	8,12	9,03	9,78	10,50	11,10	11,70	12,20	12,70
9	4,03	5,34	6,31	7,11	7,80	8,41	8,95	9,45	9,91	10,30	10,70
10	3,72	4,85	5,67	6,34	6,92	7,42	7,87	8,28	8,66	9,01	9,34
12	3,28	4,16	4,79	5,30	5,72	6,09	6,42	6,72	7,00	7,25	7,48
15	2,86	3,54	4,01	4,37	4,68	4,95	5,19	5,40	5,59	5,77	5,93
20	2,46	2,95	3,29	3,54	3,76	3,94	4,10	4,24	4,37	4,49	4,59
30	2,07	2,40	2,61	2,78	2,91	3,02	3,12	3,21	3,29	3,36	3,39
60	1,67	1,85	1,96	2,04	2,11	2,17	2,22	2,26	2,30	2,33	2,36
∞	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Tabela A10. Tabela de F máximo ou H de Hartley para $\alpha = 0,01$.

	Número de estimativas de variâncias ($\alpha = 0,01$). GL = $J-1$										
GL	2	3	4	5	6	7	8	9	10	11	12
3	47,50	85,0	120,0	151,0	184,0	216,0	249,0	281,0	310,0	337,0	361,0
4	23,20	37,0	49,0	59,0	69,0	79,0	89,0	97,0	106,0	113,0	120,0
5	14,90	22,0	28,0	33,0	38,0	42,0	46,0	50,0	54,0	57,0	60,0
6	11,10	15,5	19,1	22,0	25,0	27,0	30,0	32,0	34,0	36,0	37,0
7	8,89	12,1	14,5	16,5	18,4	20,0	22,0	23,0	24,0	26,0	27,0
8	7,50	9,9	11,7	13,2	14,5	15,8	16,9	17,9	18,9	19,8	21,0
9	6,54	8,5	9,9	11,1	12,1	13,1	13,9	14,7	15,3	16,0	16,6
10	5,85	7,4	8,6	9,6	10,4	11,1	11,8	12,4	12,9	13,4	13,9
12	4,91	6,1	6,9	7,6	8,2	8,7	9,1	9,5	9,9	10,2	10,6
15	4,07	4,9	5,5	6,0	6,4	6,7	7,1	7,3	7,5	7,8	8,0
20	3,32	3,8	4,3	4,6	4,9	5,1	5,3	5,5	5,6	5,8	5,9
30	2,63	3,0	3,3	3,4	3,6	3,7	3,8	3,9	4,0	4,1	4,2
60	1,96	2,2	2,3	2,4	2,4	2,5	2,5	2,6	2,6	2,7	2,7
∞	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Tabela A11. Limites superiores para o teste de Lilliefors.

n	$\alpha = 0,20$	$\alpha = 0,15$	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
4	0,300	0,319	0,352	0,381	0,417
5	0,285	0,299	0,315	0,337	0,405
6	0,265	0,277	0,294	0,319	0,364
7	0,247	0,258	0,276	0,300	0,348
8	0,233	0,244	0,261	0,285	0,331
9	0,223	0,233	0,249	0,271	0,311
10	0,215	0,224	0,239	0,258	0,294
11	0,206	0,217	0,230	0,249	0,284
12	0,199	0,212	0,223	0,242	0,275
13	0,190	0,202	0,214	0,234	0,268
14	0,183	0,194	0,207	0,227	0,261
15	0,177	0,187	0,201	0,220	0,257
16	0,173	0,182	0,195	0,213	0,250
17	0,169	0,177	0,189	0,206	0,245
18	0,166	0,173	0,184	0,200	0,239
19	0,163	0,169	0,179	0,195	0,235
20	0,160	0,166	0,174	0,190	0,231
25	0,142	0,147	0,158	0,173	0,200
30	0,131	0,136	0,144	0,161	0,187
> 30	$0,736/\sqrt{n}$	$0,768/\sqrt{n}$	$0,805/\sqrt{n}$	$0,886/\sqrt{n}$	$1,031/\sqrt{n}$

Tabela A12. Valores de “t” (bilaterais) para o teste de Bonferroni para $\alpha = 0,05$.

	Número de contrastes, preferencialmente ortogonais, escolhidos previamente.										
GL	2	3	4	5	6	7	8	9	10	15	20
5	3,17	3,54	3,81	4,04	4,22	4,38	4,53	4,66	4,78	5,25	5,60
7	2,84	3,13	3,34	3,50	3,64	3,76	3,86	3,95	4,03	4,36	4,59
10	2,64	2,87	3,04	3,17	3,28	3,37	3,45	3,52	3,58	3,83	4,01
12	2,56	2,78	2,94	3,06	3,15	3,24	3,31	3,37	3,43	3,65	3,80
15	2,49	2,69	2,84	2,95	3,04	3,11	3,18	3,24	3,29	3,48	3,62
20	2,42	2,61	2,75	2,85	2,93	3,00	3,06	3,11	3,16	3,33	3,46
24	2,39	2,58	2,70	2,80	2,88	2,94	3,00	3,05	3,09	3,26	3,38
30	2,36	2,54	2,66	2,75	2,83	2,89	2,94	2,99	3,03	3,19	3,30
40	2,33	2,50	2,62	2,71	2,78	2,84	2,89	2,93	2,97	3,12	3,23
60	2,30	2,47	2,58	2,66	2,73	2,79	2,84	2,88	2,92	3,06	3,16
120	2,27	2,43	2,54	2,62	2,68	2,74	2,79	2,83	2,86	2,99	3,09
∞	2,24	2,39	2,50	2,58	2,64	2,69	2,74	2,77	2,81	2,94	3,02

Tabela A13. Valores de “*t*” (bilaterais) para o teste de Bonferroni para $\alpha = 0,01$.

	Número de contrastes, preferencialmente ortogonais, escolhidos previamente.										
GL	2	3	4	5	6	7	8	9	10	15	20
5	4,78	5,25	5,60	5,89	6,15	6,36	6,56	6,70	6,86	7,51	8,00
7	4,03	4,36	4,59	4,78	4,95	5,09	5,21	5,31	5,40	5,79	6,08
10	3,58	3,83	4,01	4,15	4,27	4,37	4,45	4,53	4,59	4,86	5,06
12	3,43	3,65	3,80	3,93	4,04	4,13	4,20	4,26	4,32	4,56	4,73
15	3,29	3,48	3,62	3,74	3,82	3,90	3,97	4,02	4,07	4,29	4,42
20	3,16	3,33	3,46	3,55	3,63	3,70	3,76	3,80	3,85	4,03	4,15
24	3,09	3,26	3,38	3,47	3,54	3,61	3,66	3,70	3,74	3,91	4,04
30	3,03	3,19	3,30	3,39	3,46	3,52	3,57	3,61	3,65	3,80	3,90
40	2,97	3,12	3,23	3,31	3,38	3,43	3,48	3,51	3,55	3,70	3,79
60	2,92	3,06	3,16	3,24	3,30	3,34	3,39	3,42	3,46	3,59	3,69
120	2,86	2,99	3,09	3,16	3,22	3,27	3,31	3,34	3,37	3,50	3,58
∞	2,81	2,94	3,02	3,09	3,15	3,19	3,23	3,26	3,29	3,40	3,48

Tabela A14. Teste de Tukey: valores da amplitude total estudentizada $\alpha = 0,05$.

	Número de tratamentos								
GL	2	3	4	5	6	7	8	9	10
1	17,97	26,98	32,82	37,08	40,41	43,12	45,40	47,36	49,70
2	6,09	8,33	9,80	10,88	11,74	12,44	13,03	13,54	13,99
3	4,50	5,91	6,83	7,50	8,04	8,48	8,85	9,18	9,46
4	3,93	5,04	5,76	6,29	6,71	7,05	7,35	7,60	7,83
5	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	7,00
6	3,46	4,34	4,90	5,31	5,63	5,90	6,12	6,32	6,49
7	3,34	4,17	4,68	5,06	5,36	5,61	5,82	6,00	6,16
8	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92
9	3,20	3,95	4,42	4,76	5,02	5,24	5,43	5,60	5,74
10	3,15	3,88	4,33	4,65	4,91	5,12	5,31	5,46	5,60
11	3,11	3,82	4,26	4,57	4,82	5,03	5,20	5,35	5,49
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,40
13	3,06	3,74	4,15	4,45	4,69	4,89	5,05	5,19	5,32
14	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25
15	3,01	3,67	4,08	4,37	4,60	4,78	4,94	5,08	5,20
16	3,00	3,65	4,05	4,33	4,56	4,74	4,90	5,03	5,15
17	2,98	3,63	4,02	4,30	4,52	4,71	4,86	4,99	5,11
18	2,97	3,61	4,00	4,28	4,50	4,67	4,82	4,96	5,07
19	2,96	3,59	3,98	4,25	4,47	4,65	4,79	4,92	5,04
20	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01
24	2,92	3,53	3,90	4,17	4,37	4,54	4,68	4,81	4,92
30	2,89	3,49	3,85	4,10	4,30	4,46	4,60	4,72	4,82
40	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,64	4,74
60	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65
120	2,80	3,36	3,69	3,92	4,10	4,24	4,36	4,47	4,56
∞	2,77	3,31	3,63	3,86	4,03	4,17	4,29	4,39	4,47

Tabela A14. Teste de Tukey: amplitude total estudentizada $\alpha = 0,05$ (Continuação).

GL	Número de tratamentos									
	11	12	13	14	15	16	17	18	19	20
1	50,59	51,96	53,20	54,33	55,36	56,32	57,22	58,04	58,83	59,56
2	14,39	14,75	15,08	15,38	15,65	15,91	16,14	16,37	16,57	16,77
3	9,72	9,95	10,15	10,35	10,53	10,69	10,84	10,98	11,11	11,24
4	8,03	8,21	8,37	8,53	8,66	8,79	8,91	9,03	9,13	9,23
5	7,17	7,32	7,47	7,60	7,72	7,83	7,93	8,03	8,12	8,21
6	6,65	6,79	6,92	7,03	7,14	7,24	7,34	7,43	7,51	7,59
7	6,30	6,43	6,55	6,66	6,76	6,85	6,94	7,02	7,10	7,17
8	6,05	6,18	6,29	6,39	6,48	6,57	6,65	6,73	6,80	6,87
9	5,87	5,98	6,09	6,19	6,28	6,36	6,44	6,51	6,58	6,64
10	7,72	5,83	5,94	6,03	6,11	6,19	6,27	6,34	6,41	6,47
11	5,61	5,71	5,81	5,90	5,98	6,06	6,13	6,20	6,27	6,33
12	5,51	5,62	5,71	5,80	5,88	5,95	6,02	6,09	6,15	6,21
13	5,43	5,53	5,63	5,71	5,79	5,86	5,93	6,00	6,06	6,11
14	5,36	5,46	5,55	5,64	5,71	5,79	5,85	5,92	5,97	6,03
15	5,31	5,40	5,49	5,57	5,65	5,72	5,79	5,85	5,90	5,96
16	5,26	5,35	5,44	5,52	5,59	5,66	5,73	5,79	5,84	5,90
17	5,21	5,31	5,39	5,47	5,54	5,61	5,68	5,73	5,79	5,84
18	5,17	5,27	5,35	5,43	5,50	5,57	5,63	5,69	5,74	5,79
19	5,14	5,23	5,32	5,39	5,46	5,53	5,59	5,65	5,70	5,75
20	5,11	5,20	5,28	5,36	5,43	5,49	5,55	5,61	5,66	5,71
24	5,01	5,10	5,18	5,25	5,32	5,38	5,44	5,49	5,55	5,59
30	4,92	5,00	5,08	5,15	5,21	5,27	5,33	5,38	5,43	5,48
40	4,82	4,90	4,98	5,04	5,11	5,16	5,22	5,27	5,31	5,36
60	4,73	4,81	4,88	4,94	5,00	5,06	5,11	5,15	5,20	5,24
120	4,64	4,71	4,78	4,84	4,90	4,95	5,00	5,04	5,09	5,13
∞	4,55	4,62	4,69	4,74	4,80	4,85	4,89	4,93	4,97	5,01

Tabela A15. Teste de Tukey: valores da amplitude total estudentizada $\alpha = 0,01$.

GL	Número de tratamentos								
	2	3	4	5	6	7	8	9	10
1	90,03	135,0	164,3	185,6	202,2	215,8	227,2	237,0	245,6
2	14,04	19,02	22,29	24,72	26,63	28,20	29,53	30,68	31,69
3	8,26	10,62	12,17	13,33	14,24	15,00	15,64	16,20	16,69
4	6,51	8,12	9,17	9,96	10,58	11,10	11,55	11,93	12,27
5	5,70	6,98	7,80	8,42	8,91	9,32	9,70	9,97	10,24
6	5,24	6,33	7,03	7,56	7,97	8,32	8,61	8,87	9,10
7	4,95	5,92	6,54	7,00	7,37	7,68	7,94	8,17	8,37
8	4,75	5,64	6,20	6,62	6,96	7,24	7,47	7,68	7,86
9	4,60	5,43	5,96	6,35	6,66	6,92	7,13	7,32	7,50
10	4,48	5,27	5,77	6,14	6,43	6,67	6,88	7,00	7,21
11	4,39	5,15	5,62	5,97	6,25	6,48	6,67	6,84	6,99
12	4,32	5,05	5,50	5,84	6,10	6,32	6,51	6,67	6,81
13	4,26	4,96	5,40	5,73	5,98	6,19	6,37	6,53	6,67
14	4,21	4,90	5,32	5,63	5,88	6,08	6,26	6,41	6,54
15	4,17	4,84	5,25	5,56	5,80	5,99	6,16	6,31	6,44
16	4,13	4,79	5,19	5,49	5,72	5,92	6,08	6,22	6,35
17	4,10	4,74	5,14	5,43	5,66	5,85	6,01	6,15	6,27
18	4,07	4,70	5,09	5,38	5,60	5,79	5,94	6,08	6,20
19	4,05	4,67	5,05	5,33	5,55	5,74	5,89	6,02	6,14
20	4,02	4,64	5,02	5,29	5,51	5,69	5,84	5,97	6,09
24	3,96	4,55	4,91	5,17	5,37	5,54	5,68	5,81	5,92
30	3,89	4,46	4,80	5,05	5,24	5,40	5,54	5,65	5,76
40	3,82	4,37	4,70	4,93	5,11	5,26	5,39	5,50	5,60
60	3,76	4,28	4,60	4,82	4,99	5,13	5,25	5,36	5,45
120	3,70	4,20	4,50	4,71	4,87	5,00	5,12	5,21	5,30
∞	3,64	4,12	4,40	4,60	4,76	4,88	4,99	5,08	5,16

Tabela A15. Teste de Tukey: amplitude total estudentizada $\alpha = 0,01$ (Continuação).

GL	Número de tratamentos									
	11	12	13	14	15	16	17	18	19	20
1	253,2	260,0	266,2	271,8	277,0	281,8	286,3	290,4	294,3	298,0
2	32,59	33,40	34,13	34,81	35,43	36,00	36,53	37,03	37,50	37,95
3	17,13	17,53	17,89	18,22	18,52	18,81	19,07	19,32	19,55	19,77
4	12,57	12,84	13,09	13,32	13,53	13,73	13,91	14,08	14,24	14,40
5	10,48	10,70	10,89	11,08	11,24	11,40	11,55	11,68	11,81	11,93
6	9,30	9,48	9,65	9,81	9,95	10,08	10,21	10,32	10,43	10,54
7	8,55	8,71	8,86	9,00	9,12	9,24	9,35	9,46	9,55	9,65
8	8,03	8,18	8,31	8,44	8,55	8,66	8,76	8,85	8,94	9,03
9	7,65	7,78	7,91	8,02	8,13	8,23	8,32	8,41	8,50	8,58
10	7,36	7,48	7,60	7,71	7,81	7,91	7,99	8,08	8,15	8,23
11	7,13	7,25	7,36	7,46	7,56	7,65	7,73	7,81	7,88	7,95
12	6,94	7,06	7,17	7,26	7,36	7,44	7,52	7,59	7,66	7,73
13	6,79	6,90	7,01	7,10	7,19	7,27	7,34	7,42	7,43	7,55
14	6,66	6,77	6,87	6,96	7,05	7,13	7,20	7,27	7,33	7,40
15	6,56	6,66	6,76	6,84	6,93	7,00	7,07	7,14	7,20	7,26
16	6,46	6,56	6,66	6,74	6,82	6,90	6,97	7,03	7,09	7,15
17	6,38	6,48	6,57	6,66	6,73	6,81	6,87	6,94	7,00	7,05
18	6,31	6,41	6,50	6,58	6,66	6,72	6,79	6,85	6,91	6,97
19	6,25	6,34	6,43	6,51	6,58	6,65	6,72	6,78	6,84	6,89
20	6,19	6,28	6,37	6,45	6,52	6,59	6,65	6,71	6,77	6,82
24	6,02	6,11	6,19	6,26	6,33	6,39	6,45	6,51	6,56	6,61
30	5,85	5,93	6,01	6,08	6,14	6,20	6,26	6,31	6,36	6,41
40	5,69	5,76	5,84	5,90	5,96	6,02	6,07	6,12	6,16	6,21
60	5,53	5,60	5,67	5,73	5,78	5,84	5,89	5,93	5,97	6,02
120	5,38	5,44	5,50	5,56	5,61	5,66	5,71	5,75	5,79	5,83
∞	5,23	5,29	5,35	5,40	5,45	5,49	5,54	5,57	5,61	5,64

Tabela A16. Teste de Duncan: valores da amplitude total estudentizada $\alpha = 0,05$.

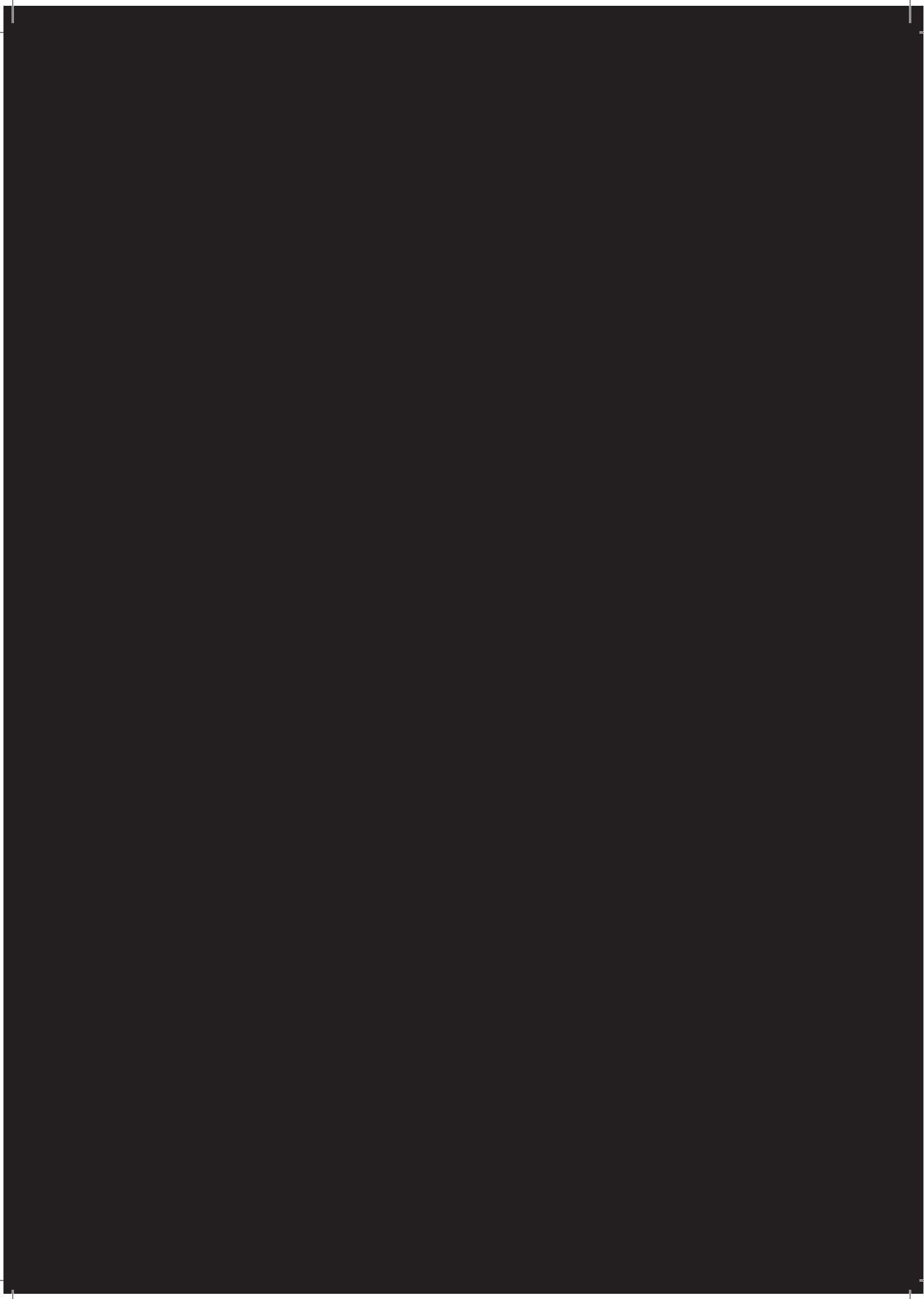
GL	Número de tratamentos												
	2	3	4	5	6	7	8	9	10	12	14	16	18
1	18,0	18,0	18,0	18,0	18,0	18,0	18,0	18,0	18,0	18,0	18,0	18,0	18,0
2	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09	6,09
3	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50	4,50
4	3,93	4,01	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02	4,02
5	3,64	3,74	3,79	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83	3,83
6	3,46	3,58	3,64	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68	3,68
7	3,35	3,47	3,54	3,58	3,60	3,61	3,61	3,61	3,61	3,61	3,61	3,61	3,61
8	3,26	3,39	3,47	3,52	3,55	3,56	3,56	3,56	3,56	3,56	3,56	3,56	3,56
9	3,20	3,34	3,41	3,47	3,50	3,52	3,52	3,52	3,52	3,52	3,52	3,52	3,52
10	3,15	3,30	3,37	3,43	3,46	3,47	3,47	3,47	3,47	3,47	3,47	3,47	3,47
11	3,11	3,27	3,35	3,39	3,43	3,44	3,45	3,46	3,46	3,46	3,46	3,46	3,47
12	3,08	3,23	3,33	3,36	3,40	3,42	3,44	3,44	3,46	3,46	3,46	3,46	3,47
13	3,06	3,21	3,30	3,35	3,38	3,41	3,42	3,44	3,45	3,45	3,46	3,46	3,47
14	3,03	3,18	3,27	3,33	3,37	3,39	3,41	3,42	3,44	3,45	3,46	3,46	3,47
15	3,01	3,16	3,25	3,31	3,36	3,38	3,40	3,42	3,43	3,44	3,45	3,46	3,47
16	3,00	3,15	3,23	3,30	3,34	3,37	3,39	3,41	3,43	3,44	3,45	3,46	3,47
17	2,98	3,13	3,22	3,28	3,33	3,36	3,38	3,40	3,42	3,44	3,45	3,46	3,47
18	2,97	3,12	3,21	3,27	3,32	3,35	3,37	3,39	3,41	3,43	3,45	3,46	3,47
19	2,96	3,11	3,19	3,26	3,31	3,35	3,37	3,39	3,41	3,43	3,44	3,46	3,47
20	2,95	3,10	3,18	3,25	3,30	3,34	3,36	3,38	3,40	3,43	3,44	3,46	3,46
22	2,93	3,08	3,17	3,24	3,29	3,32	3,35	3,37	3,39	3,42	3,44	3,45	3,46
24	2,92	3,07	3,15	3,22	3,28	3,31	3,34	3,37	3,38	3,41	3,44	3,45	3,46
26	2,91	3,06	3,14	3,21	3,27	3,30	3,34	3,36	3,38	3,41	3,43	3,45	3,46
28	2,90	3,04	3,13	3,20	3,26	3,30	3,33	3,35	3,37	3,40	3,43	3,45	3,46
30	2,89	3,04	3,12	3,20	3,25	3,29	3,32	3,35	3,37	3,40	3,43	3,44	3,46
40	2,86	3,01	3,10	3,17	3,22	3,27	3,30	3,33	3,35	3,39	3,42	3,44	3,46
60	2,83	2,98	3,08	3,14	3,20	3,24	3,28	3,31	3,33	3,37	3,40	3,43	3,45
100	2,80	2,95	3,05	3,12	3,18	3,22	3,26	3,29	3,32	3,36	3,40	3,42	3,45
∞	2,77	2,92	3,02	3,09	3,15	3,19	3,23	3,26	3,29	3,34	3,38	3,41	3,44

Tabela A17. Teste de Duncan: valores da amplitude total estudentizada $\alpha = 0,01$.

GL	Número de tratamentos												
	2	3	4	5	6	7	8	9	10	12	14	16	18
1	90,0	90,0	90,0	90,0	90,0	90,0	90,0	90,0	90,0	90,0	90,0	90,0	90,0
2	14,0	14,0	14,0	14,0	14,0	14,0	14,0	14,0	14,0	14,0	14,0	14,0	14,0
3	8,26	8,50	8,60	8,70	8,80	8,90	8,90	9,00	9,00	9,00	9,10	9,20	9,30
4	6,51	6,80	6,90	7,00	7,10	7,10	7,20	7,20	7,30	7,30	7,40	7,40	7,50
5	5,70	5,96	6,11	6,18	6,26	6,33	6,40	6,44	6,50	6,60	6,60	6,70	6,70
6	5,24	5,51	5,65	5,73	5,81	5,88	5,95	6,00	6,00	6,10	6,20	6,20	6,30
7	4,95	5,22	5,37	5,45	5,53	5,61	5,69	5,73	5,80	5,80	5,90	5,90	6,00
8	4,74	5,00	5,14	5,23	5,32	5,40	5,47	5,51	5,50	5,60	5,70	5,70	5,80
9	4,60	4,86	4,99	5,08	5,17	5,25	5,32	5,36	5,40	5,50	5,50	5,60	5,70
10	4,48	4,73	4,88	4,96	5,06	5,13	5,20	5,24	5,28	5,36	5,42	5,48	5,54
11	4,39	4,63	4,77	4,86	4,94	5,01	5,06	5,12	5,15	5,24	5,28	5,34	5,38
12	4,32	4,55	4,68	4,76	4,84	4,92	4,96	5,02	5,07	5,13	5,17	5,22	5,24
13	4,26	4,48	4,62	4,69	4,74	4,84	4,88	4,94	4,98	5,04	5,08	5,13	5,14
14	4,21	4,42	4,55	4,63	4,70	4,78	4,83	4,87	4,91	4,96	5,00	5,04	5,06
15	4,17	4,37	4,50	4,58	4,64	4,72	4,77	4,81	4,84	4,90	4,94	4,97	4,99
16	4,13	4,34	4,45	4,54	4,60	4,67	4,72	4,76	4,79	4,84	4,88	4,91	4,93
17	4,10	4,30	4,41	4,50	4,56	4,63	4,68	4,72	4,75	4,80	4,83	4,86	4,88
18	4,07	4,27	4,38	4,46	4,53	4,59	4,64	4,68	4,71	4,76	4,79	4,82	4,84
19	4,05	4,24	4,35	4,43	4,50	4,56	4,61	4,64	4,67	4,72	4,76	4,79	4,81
20	4,02	4,22	4,33	4,40	4,47	4,53	4,58	4,61	4,65	4,69	4,73	4,76	4,78
22	3,99	4,17	4,28	4,36	4,42	4,48	4,53	4,57	4,60	4,65	4,68	4,71	4,74
24	3,96	4,14	4,24	4,33	4,39	4,44	4,49	4,53	4,57	4,62	4,64	4,67	4,70
26	3,93	4,11	4,21	4,30	4,36	4,41	4,46	4,50	4,53	4,58	4,62	4,65	4,67
28	3,91	4,08	4,18	4,28	4,34	4,39	4,43	4,47	4,51	4,56	4,60	4,62	4,65
30	3,89	4,06	4,16	4,22	4,32	4,36	4,41	4,45	4,48	4,54	4,58	4,61	4,63
40	3,82	3,99	4,10	4,17	4,24	4,30	4,34	4,37	4,41	4,46	4,51	4,54	4,57
60	3,76	3,92	4,03	4,12	4,17	4,23	4,27	4,31	4,34	4,39	4,44	4,47	4,50
100	3,71	3,86	3,98	4,06	4,11	4,17	4,21	4,25	4,29	4,35	4,38	4,42	4,45
∞	3,64	3,80	3,90	3,98	4,04	4,09	4,14	4,17	4,20	4,26	4,31	4,34	4,38

Tabela A18. Quantis de ordem ($p = 1 - \alpha$) para o teste de Shapiro-Wilk (W_p).

n	$W_{0,01}$	$W_{0,02}$	$W_{0,05}$	$W_{0,10}$	$W_{0,50}$	$W_{0,90}$	$W_{0,95}$
3	0,753	0,756	0,767	0,789	0,959	0,998	0,999
4	0,687	0,707	0,748	0,792	0,935	0,987	0,992
5	0,686	0,715	0,762	0,803	0,927	0,979	0,986
6	0,713	0,743	0,788	0,826	0,927	0,974	0,981
7	0,730	0,760	0,803	0,838	0,928	0,972	0,979
8	0,749	0,778	0,818	0,851	0,932	0,972	0,978
9	0,764	0,791	0,829	0,859	0,935	0,972	0,978
10	0,781	0,806	0,842	0,869	0,938	0,972	0,978
11	0,792	0,817	0,850	0,876	0,940	0,973	0,979
12	0,805	0,828	0,859	0,883	0,943	0,973	0,979
13	0,814	0,837	0,866	0,889	0,945	0,974	0,979
14	0,825	0,846	0,874	0,895	0,947	0,975	0,980
15	0,835	0,855	0,881	0,901	0,950	0,975	0,980
16	0,844	0,863	0,887	0,906	0,952	0,976	0,981
17	0,851	0,869	0,892	0,910	0,954	0,977	0,981
18	0,858	0,874	0,897	0,914	0,956	0,978	0,982
19	0,863	0,879	0,901	0,917	0,957	0,978	0,982
20	0,868	0,884	0,905	0,920	0,959	0,979	0,983



APÊNDICE B

PROGRAMAS SAS

- Prog1.** Valores de $F(X)$ para o teste de Komogorov-Smirnov. Exercício 1.3.5.1.
- Prog2.** Valores de $F(X)$ para o teste de Lilliefords. Exercício 1.3.6.1.
- Prog3.** Valores de $F(X)$ para os testes de Cramér-Von Mises e Anderson-Darling. Exercícios: 1.3.7.1 e 1.3.8.1.
- Prog4.** Testes de normalidade. Exercícios 1.3.7.1, 1.3.8.1 e 1.3.9.1.
- Prog5a.** Testes de normalidade do ensaio casualizado em blocos. Exercício 4.2.1.
- Prog5b.** Testes de normalidade do ensaio casualizado em blocos. Exercício 4.2.1.
- Prog6.** Testes de homogeneidade de variâncias. Exercício 2.5.1.
- Prog7.** Teste de Tukey de não aditividade de efeitos no DCB. Exercício 4.2.1.
- Prog8.** Experimento inteiramente ao acaso com K amostras por parcela. Exercício 2.6.1.
- Prog9.** Análise de covariância no ensaio inteiramente ao acaso. Exercício 2.7.1.1.
- Prog10.** Teste de Tukey. Exercício 3.2.1.
- Prog11.** Teste de Tukey. Exercício 3.2.2.
- Prog12.** Teste de Sheffé. Exercício 3.3.1.
- Prog13.** Teste de Duncan. Exercício 3.4.1.
- Prog14.** Teste de Dunnett. Exercício 3.6.1.
- Prog15.** Teste de Bonferroni. Exercício 3.7.1.
- Prog16.** Análise de ensaio em blocos com tratamentos comuns. Exercício 4.4.1.1.
- Prog17.** Análise de regressão pela equação de Mitscherlich. Exercício 4.5.2.1.
- Prog18.** Análise em quadrado latino. Exercício 5.3.1.
- Prog19.** Análise ensaio fatorial com dois fatores. Exercício 6.2.1.
- Prog20.** Experimento fatorial hierárquico em dois estágios. Exercício 6.3.1.1
- Prog21.** Experimento fracionário 2^{4-1}_{III} . Exercício 6.5.4.1.
- Prog22.** Análise de grupos de experimentos. Exercício 7.3.1.
- Prog23.** Análise de grupos de experimentos. Exercício 7.3.2.
- Prog24.** Análise de experimento em parcelas subdivididas. Exercício 8.1.1.
- Prog25.** Análise de experimento em faixas. Exercício 8.2.1.
- Prog26.** Análise BIB. Tipo I. Análise intrablocos. Exercício 9.3.1.1.
- Prog27.** Análise BIB. Tipo II. Análise intrablocos. Exercício 9.3.1.2
- Prog28.** Análise BIB. Tipo III. Análise intrablocos. Exercício 9.3.1.3
- Prog29.** Látice quadrado (*Square Lattice*). Análise intrablocos. Exercício 10.2.1.1.
- Prog30.** Látice quadrado (*Square Lattice*). Análise intrablocos. Exercício 10.2.1.2.
- Prog31.** Látice quadrado (*Square Lattice*). Análise com recuperação da informação interblocos. Exercício 10.2.2.1.
- Prog32.** Látice quadrado (*Square Lattice*). Análise com recuperação da informação interblocos. Exercício 10.2.2.2.
- Prog33.** Superfície de resposta fatorial 3×3 . Exercício 11.1.5.1.
- Prog34.** Operações com matrizes. Fatorial 3×3 . Exercício 11.1.5.1.
- Prog35.** Superfície de resposta fatorial $3 \times 3 \times 3$. Exercício 11.2.3.1
- Prog36.** Superfície de resposta. Delineamento composto central. Exercício 11.3.1.1.
- Prog37.** Operações com matrizes. Delineamento composto central. Exercício 11.3.1.1
- Prog38.** Delineamento Box-Behnken. Exercício 11.4.1.
- Prog39.** Aplicação de um fatorial $(\frac{1}{5})5^3$. Exercício 11.5.1.

Prog1. Valores de para o teste de Komogorov-Smirnov. Exercício 1.3.5.1.

```
proc iml;
start fun(x);
v = (1/4.2612680668727)*exp(-0.5*((x-13.3)/1.7)**2);
return(v);
finish;
x={0 11.0};call quad(z,"fun",x);print z;
x={0 11.3};call quad(z,"fun",x);print z;
x={0 12.1};call quad(z,"fun",x);print z;
x={0 13.3};call quad(z,"fun",x);print z;
x={0 13.6};call quad(z,"fun",x);print z;
x={0 13.7};call quad(z,"fun",x);print z;
x={0 13.8};call quad(z,"fun",x);print z;
x={0 14.1};call quad(z,"fun",x);print z;
x={0 14.2};call quad(z,"fun",x);print z;
x={0 14.7};call quad(z,"fun",x);print z;
x={0 14.8};call quad(z,"fun",x);print z;
x={0 15.4};call quad(z,"fun",x);print z;
x={0 15.8};call quad(z,"fun",x);print z;
x={0 16.0};call quad(z,"fun",x);print z;
x={0 16.2};call quad(z,"fun",x);print z; quit;
```

Prog2 Valores de para o teste de Lilliefords. Exercício 1.3.6.1.proc iml;

```
proc iml;
start fun(x);
v = (1/4.0137373017685)*exp(-0.5*((x-14)/1.6012495120998)**2);
return(v);
finish;
x={0 11.0};call quad(z,"fun",x);print z;
x={0 11.3};call quad(z,"fun",x);print z;
x={0 12.1};call quad(z,"fun",x);print z;
x={0 13.3};call quad(z,"fun",x);print z;
x={0 13.6};call quad(z,"fun",x);print z;
x={0 13.7};call quad(z,"fun",x);print z;
x={0 13.8};call quad(z,"fun",x);print z;
x={0 14.1};call quad(z,"fun",x);print z;
x={0 14.2};call quad(z,"fun",x);print z;
x={0 14.7};call quad(z,"fun",x);print z;
x={0 14.8};call quad(z,"fun",x);print z;
x={0 15.4};call quad(z,"fun",x);print z;
x={0 15.8};call quad(z,"fun",x);print z;
x={0 16.0};call quad(z,"fun",x);print z;
x={0 16.2};call quad(z,"fun",x);print z;
quit;
```

Prog3. Valores de para os testes de Cramér-Von Mises e Anderson-Darling. Exercícios: 1.3.7.1 e 1.3.8.1.

```
proc iml;
start fun(x);
v = (1/2.0821618334465)*exp(-0.5*((x-12.7)/0.83066239)**2);
return(v);
finish;
x={0 11.8};call quad(z,"fun",x);print z;
x={0 11.9};call quad(z,"fun",x);print z;
x={0 12.8};call quad(z,"fun",x);print z;
x={0 13.4};call quad(z,"fun",x);print z;
x={0 13.6};call quad(z,"fun",x);print z;
quit;
```

Prog4. Testes de normalidade. Exercícios 1.3.7.1, 1.3.8.1 e 1.3.9.1.

```
data ARV;
input alt;
cards;
11.8
11.9
12.8
13.4
13.6
```

```
;
ODS HTML FILE ="ARV";
ODS GRAPHICS ON;
PROC UNIVARIATE NORMAL PLOT;
VAR alt;
run;
ods html close;
run;
quit;
```

Prog5a. Testes de normalidade. Ensaio casualizado em blocos. Exercício 4.2.1.

```
data DBC;
input Bloco Trat $ Prod@@;
cards;
1 E1 43.3 2 E1 40.4 3 E1 42.5 4 E1 38.5 5 E1 44.3
1 E2 59.7 2 E2 52.8 3 E2 63.0 4 E2 51.3 5 E2 56.3
1 E3 56.7 2 E3 56.7 3 E3 60.5 4 E3 70.0 5 E3 66.2
1 E4 49.0 2 E4 46.7 3 E4 54.0 4 E4 46.4 5 E4 51.8
;
ods html file='DBC';
ods graphics on;
proc glm data=DBC;
class Bloco Trat;
model Prod = Bloco Trat;
output out=teste P=predito R=residuo;
run;
proc univariate data=teste normal;
var residuo;
run;
ods html close;
run;
quit;
```

Prog5b. Testes de normalidade. Ensaio casualizado em blocos. Exercício 4.2.1.

```
data ARV;
input alt@@;
cards;
1.83 1.955 -1.795 -2.345 0.355 3.41 -0.465 3.885 -4.365 -2.465
-4.99 -1.965 -4.015 8.935 2.035 -0.25 0.475 1.925 -2.225 0.075
;
ODS HTML FILE ="ARV";
ODS GRAPHICS ON;
PROC UNIVARIATE NORMAL PLOT;
VAR alt;
run;
ods html close;
run;
quit;
```

Prog6. Testes de homogeneidade de variâncias. Exercício 2.5.1.

```
data dic;
input trat $ nim @@;
cards;
A 5546 A 5406 A 5233 A 5646 A 5412 A 5623
B 1648 B 1692 B 1845 B 1810 B 1536 B 1541
C 926 C 892 C 857 C 801 C 667 C 773
D 208 D 198 D 268 D 189 D 167 D 134
E 158 E 189 E 207 E 142 E 172 E 168
;
ods html file='dic';
ods graphics on;
proc anova;
class trat;
model nim=trat;
means trat/hovtest= bartlett;
means trat/hovtest= levene;
means trat/hovtest= obrien;
run;
ods html close;
run;
quit;
```

Prog7. Teste de não aditividade de Tukey: ensaio casualizado em blocos (DCB). Exercício 4.2.1.

```

data DBC;
input Bloco Trat $ Prod@@;
cards;
1 E1 43.3 2 E1 40.4 3 E1 42.5 4 E1 38.5 5 E1 44.3
1 E2 59.7 2 E2 52.8 3 E2 63.0 4 E2 51.3 5 E2 56.3
1 E3 56.7 2 E3 56.7 3 E3 60.5 4 E3 70.0 5 E3 66.2
1 E4 49.0 2 E4 46.7 3 E4 54.0 4 E4 46.4 5 E4 51.8
;
ods html file="DBC";
ods graphics on;
proc glm data=DBC;
class Bloco Trat;
model Prod = Bloco Trat;
output out=teste P=vp1 R=residuo;
run;
data teste; set teste;
VP=vp1*vp1;
run;
proc glm data=teste;
class Bloco Trat;
model Prod= Bloco Trat VP;
run;
ods html close;
run;
quit;

```

Prog8. Experimento inteiramente ao acaso com K amostras por parcela. Exercício 2.6.1.

```

data DICCS;
input TRAT $ VASO PLANTA ALT@@;
cards;
E1 1 1 8.2 E1 1 2 9.3 E1 1 3 7.0 E1 1 4 10.5
E1 2 1 5.8 E1 2 2 10.4 E1 2 3 12.9 E1 2 4 11.7
E1 3 1 12.0 E1 3 2 7.0 E1 3 3 5.8 E1 3 4 7.0
E2 4 1 11.7 E2 4 2 12.0 E2 4 3 9.3 E2 4 4 8.2
E2 5 1 8.2 E2 5 2 8.2 E2 5 3 7.0 E2 5 4 9.3
E2 6 1 10.5 E2 6 2 9.3 E2 6 3 9.3 E2 6 4 11.7
E3 7 1 11.7 E3 7 2 10.5 E3 7 3 15.7 E3 7 4 10.5
E3 8 1 12.9 E3 8 2 14.0 E3 8 3 11.7 E3 8 4 11.7
E3 9 1 12.9 E3 9 2 10.5 E3 9 3 15.2 E3 9 4 12.9
E4 10 1 19.9 E4 10 2 18.0 E4 10 3 21.1 E4 10 4 19.9
E4 11 1 15.2 E4 11 2 16.4 E4 11 3 18.7 E4 11 4 15.2
E4 12 1 8.5 E4 12 2 16.3 E4 12 3 12.4 E4 12 4 20.1
;
ods html file="DICCS";
ods graphics on;
proc glm data=DICCS;
class TRAT VASO;
model ALT= TRAT VASO(TRAT)/SS3;
test H=TRAT E=VASO(TRAT);
means TRAT/TuKey E=VASO(TRAT);
run;
ods html close;
run;
quit;

```

Prog9. Análise de covariância: Ensaio inteiramente ao acaso. Exercício 2.7.1.1.

```

data ANCOVA;
input TRAT $ NA VOL@@;
cards;
M1 30 66.9 M1 32 68.2 M1 35 66.0 M1 30 58.1 M1 26 57.7
M2 23 40.7 M2 25 54.5 M2 16 35.8 M2 19 28.7 M2 13 15.0
M3 27 48.7 M3 36 56.9 M3 23 48.4 M3 19 40.8 M3 10 32.4
M4 22 46.9 M4 24 49.9 M4 21 45.3 M4 23 43.8 M4 20 45.6
M5 29 55.4 M5 32 57.8 M5 21 42.7 M5 14 36.7 M5 21 46.7
;
ods html file="ANCOVA";
ods graphics on;

```

```

proc glm data=ANCOVA;
class TRAT;
model VOL= TRAT;
run;
proc glm data=ANCOVA;
class TRAT;
model NA = TRAT;
run;
proc glm data=ANCOVA;
CLASS TRAT;
MODEL VOL = TRAT NA / SOLUTION;
LSMEANS TRAT / STDERR PDIF CL ADJUST=TUKEY;
OUTPUT OUT = pred p=ybar r=resid;
RUN;
ods html close;
run;
quit;

```

Prog10. Teste de Tukey. Exercício 3.2.1.

```

data TUKEY;
input TRAT $ PROD@@;
cards;
M1 34.5 M1 38.3 M1 43.5 M1 41.3 M1 39.5 M1 44.0
M2 45.0 M2 53.3 M2 54.0 M2 51.0 M2 49.6 M2 52.3
M3 43.5 M3 54.0 M3 51.8 M3 50.3 M3 53.4 M3 45.6
M4 42.0 M4 36.8 M4 44.3 M4 42.8 M4 43.0 M4 34.1
;
ODS HTML FILE ="TUKEY";
ODS GRAPHICS ON;
PROC GLM;
CLASS TRAT;
MODEL PROD = TRAT;
MEANS TRAT/TUKEY;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog11. Teste de Tukey. Exercício 3.2.2.

```

data SHEFFE;
input TRAT $ PROD@@;
cards;
M1 34.5 M1 38.3 M1 43.5 M1 41.3 M1 39.5 M1 44.0
M2 45.0 M2 53.3 M2 54.0 M2 51.0 M2 49.6 M2 52.3
M3 43.5 M3 54.0 M3 51.8 M3 50.3 M3 53.4 M3 45.6
M4 42.0 M4 36.8 M4 44.3 M4 42.8 M4 43.0 M4 34.1
;
ODS HTML FILE ="SHEFFE";
ODS GRAPHICS ON;
PROC GLM;
CLASS TRAT;
MODEL PROD = TRAT;
MEANS TRAT/SHEFFE;
RUN;
ODS HTML CLOSE;
run;
QUIT;

```

Prog12. Teste de Sheffé. Exercício 3.3.1.

```

data SHEFFE;
input TRAT $ PROD@@;
cards;
M1 34.5 M1 38.3 M1 43.5 M1 41.3 M1 39.5 M1 44.0
M2 45.0 M2 53.3 M2 54.0 M2 51.0 M2 49.6 M2 52.3
M3 43.5 M3 54.0 M3 51.8 M3 50.3 M3 53.4 M3 45.6
M4 42.0 M4 36.8 M4 44.3 M4 42.8 M4 43.0 M4 34.1
;
ODS HTML FILE ="SHEFFE";
ODS GRAPHICS ON;
PROC GLM;

```

```

CLASS TRAT;
MODEL PROD = TRAT;
MEANS TRAT/SHEFFE;
RUN;
ODS HTML CLOSE;
run;
QUIT;

```

Prog13. Teste de Duncan. Exercício 3.4.1.

```

data DUNCAN;
input TRAT $ PROD@@;
cards;
M1 34.5 M1 38.3 M1 43.5 M1 41.3 M1 39.5 M1 44.0
M2 45.0 M2 53.3 M2 54.0 M2 51.0 M2 49.6 M2 52.3
M3 43.5 M3 54.0 M3 51.8 M3 50.3 M3 53.4 M3 45.6
M4 42.0 M4 36.8 M4 44.3 M4 42.8 M4 43.0 M4 34.1
;
ODS HTML FILE ="DUNCAN";
ODS GRAPHICS ON;
PROC GLM;
CLASS TRAT;
MODEL PROD = TRAT;
MEANS TRAT/DUNCAN;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog14. Teste de Dunnett. Exercício 3.6.1.

```

data DUNNETT;
input TRAT $ PROD@@;
cards;
M1 34.5 M1 38.3 M1 43.5 M1 41.3 M1 39.5 M1 44.0
M2 45.0 M2 53.3 M2 54.0 M2 51.0 M2 49.6 M2 52.3
M3 43.5 M3 54.0 M3 51.8 M3 50.3 M3 53.4 M3 45.6
M4 42.0 M4 36.8 M4 44.3 M4 42.8 M4 43.0 M4 34.1
;
ODS HTML FILE ="DUNNETT";
ODS GRAPHICS ON;
PROC GLM;
CLASS TRAT;
MODEL PROD = TRAT;
MEANS TRAT/DUNNETT;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog15. Teste de Bonferroni

```

data BONFERRONI;
input TRAT $ PROD@@;
cards;
M1 34.5 M1 38.3 M1 43.5 M1 41.3 M1 39.5 M1 44.0
M2 45.0 M2 53.3 M2 54.0 M2 51.0 M2 49.6 M2 52.3
M3 43.5 M3 54.0 M3 51.8 M3 50.3 M3 53.4 M3 45.6
M4 42.0 M4 36.8 M4 44.3 M4 42.8 M4 43.0 M4 34.1
;
ODS HTML FILE ="BONFERRONI";
ODS GRAPHICS ON;
PROC GLM;
CLASS TRAT;
MODEL PROD = TRAT;
MEANS TRAT/BON;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog16. Análise de ensaio em blocos com tratamentos adicionais. Exercício 4.3.1.

```

data BLOCOTAD;
input Grupos Blocos Procedencias $ Prod@@;
cards;
1 1 PC1 85.5 1 1 PC2 84.6 1 1 P1 86.3 1 1 P2 83.8 1 1 P3 87.9
1 1 P4 84.6 1 1 P5 82.2 1 1 P6 89.5 1 2 PC1 83.0 1 2 PC2 83.8
1 2 P1 87.1 1 2 P2 83.0 1 2 P3 88.7 1 2 P4 86.3 1 2 P5 83.8
1 2 P6 91.2 1 3 PC1 84.6 1 3 PC2 85.5 1 3 P1 85.5 1 3 P2 84.6
1 3 P3 91.2 1 3 P4 85.5 1 3 P5 83.0 1 3 P6 88.7 1 4 PC1 86.3
1 4 PC2 83.5 1 4 P1 86.3 1 4 P2 82.2 1 4 P3 89.5 1 4 P4 87.1
1 4 P5 82.2 1 4 P6 90.3 2 1 PC1 86.3 2 1 PC2 83.0 2 1 P7 83.8
2 1 P8 82.2 2 1 P9 84.6 2 1 P10 90.3 2 1 P11 92.0 2 1 P12 87.1
2 2 PC1 83.0 2 2 PC2 84.6 2 2 P7 86.3 2 2 P8 80.6 2 2 P9 85.5
2 2 P10 92.0 2 2 P11 90.3 2 2 P12 86.3 2 3 PC1 83.8 2 3 PC2 86.3
2 3 P9 86.3 2 3 P7 84.6 2 3 P8 81.4 2 3 P10 89.5 2 3 P11 91.2
2 3 P12 87.9 2 4 PC1 84.6 2 4 PC2 83.8 2 4 P7 82.2 2 4 P8 83.8
2 4 P9 83.0 2 4 P10 92.8 2 4 P11 93.6 2 4 P12 85.5 3 1 PC1 88.7
3 1 PC2 85.5 3 1 P13 91.2 3 1 P14 86.3 3 1 P15 83.8 3 1 P16 89.5
3 1 P17 91.2 3 1 P18 92.8 3 2 PC1 87.1 3 2 PC2 87.9 3 2 P13 91.2
3 2 P14 85.5 3 2 P15 86.3 3 2 P16 90.3 3 2 P17 93.6 3 2 P18 91.2
3 3 PC1 85.5 3 3 PC2 84.6 3 3 P13 92.0 3 3 P14 87.1 3 3 P15 87.9
3 3 P16 91.2 3 3 P17 92.8 3 3 P18 92.0 3 4 PC1 86.3 3 4 PC2 87.1
3 4 P13 89.5 3 4 P14 87.9 3 4 P15 83.0 3 4 P16 87.9 3 4 P17 94.4
3 4 P18 91.2
;
ods html file="BLOCOTAD";
ods graphics on;
proc glm data=BLOCOTAD;
class Blocos Procedencias;
model Prod= Blocos Procedencias;
BY Grupos;
run;
proc glm data=BLOCOTAD;
class Blocos Grupos Procedencias;
model Prod= Grupos Blocos(Grupos) Procedencias Grupos*Procedencias/ss1;
Run;
ods html close;
run; quit;

```

Prog17. Análise de regressão pela equação de Mitscherlich. Exercício 4.5.2.1.

```

DATA MITSCHERLICH;
INPUT Bloco Dose Prod@@;
CARDS;
1 0 63 2 0 58 3 0 69 4 0 50 5 0 49
1 30 90 2 30 105 3 30 87 4 30 71 5 30 78
1 60 111 2 60 114 3 60 112 4 60 115 5 60 104
1 90 125 2 90 110 3 90 120 4 90 115 5 90 105
1 120 115 2 120 113 3 120 120 4 120 130 5 120 121
;
ods html file="MITSCHERLICH";
ods graphics on;
PROC NLIN Method=GAUSS;
Parms b0=124 b1=0.01 b2=28;
MODEL Prod=b0*(1-10**(-b1*(Dose+b2)));
run;
PROC GPLOT;
SYMBOL Prod=DOT I=JOIN C=BLUE HEIGHT=0.5;
plot Prod*Dose;
TITLE HEIGHT=1.4 'TENDENCIA DE CRESCIMENTO';
ods html close;
run;
quit;

```

Prog18. Análise em quadrado latino. Exercício 5.3.1.

```

data DQL;
input Trat $ Linha Coluna Prod@@;
cards;
E1 1 5 45.9 E1 2 2 63.6 E1 3 1 47.2 E1 4 3 48.5 E1 5 4 58.9
E2 1 1 30.1 E2 2 5 49.5 E2 3 4 31.4 E2 4 2 45.7 E2 5 3 46.7
E3 1 4 60.0 E3 2 1 48.5 E3 3 3 58.1 E3 4 5 46.8 E3 5 2 82.4
E4 1 3 35.5 E4 2 4 46.0 E4 3 2 46.5 E4 4 1 41.2 E4 5 5 40.6

```

```
E5 1 2 59.6 E5 2 3 38.1 E5 3 5 32.1 E5 4 4 49.5 E5 5 1 31.5
```

```
;
ods html file="DQL";
ods graphics on;
proc glm data=DQL;
Class Trat Linha Coluna;
model Prod = Trat Linha Coluna;
means Trat/TuKey ;
run;
ods html close;
run;
quit;
```

Prog19. Análise ensaio fatorial com dois fatores. Exercício 6.2.1.

```
data SIMOES;
input Rec $ Esp $ Rep Alt@@;
cards;
R1 E1 1 26.2 R1 E1 2 26.0 R1 E1 3 25.0 R1 E1 4 25.4
R1 E2 1 24.8 R1 E2 2 24.6 R1 E2 3 26.7 R1 E2 4 25.2
R2 E1 1 25.7 R2 E1 2 26.3 R2 E1 3 25.1 R2 E1 4 26.4
R2 E2 1 19.6 R2 E2 2 21.1 R2 E2 3 19.0 R2 E2 4 18.6
R3 E1 1 22.8 R3 E1 2 19.4 R3 E1 3 18.8 R3 E1 4 19.2
R3 E2 1 19.8 R3 E2 2 21.4 R3 E2 3 22.8 R3 E2 4 21.3
;
ods html file="SIMOES";
ods graphics on;
proc GLM;
class Rec Esp;
model Alt = Rec Esp Rec*Esp;
run;
lsmeans Rec*Esp/slice=Esp;
lsmeans Rec*Esp/slice=Rec;
run;
ods html close;
run;
quit;
```

Prog20. Experimento fatorial hierárquico em dois estágios. Exercício 6.3.1.1

```
DATA HIER;
INPUT Especie $ Procedencia $ Rep Producao @@;
cards;
E1 P1 1 177.1 E1 P2 1 166.2 E1 P3 1 190.6 E1 P4 1 169.1
E1 P1 2 187.3 E1 P2 2 174.7 E1 P3 2 170.9 E1 P4 2 197.9
E1 P1 3 186.8 E1 P2 3 183.9 E1 P3 3 174.0 E1 P4 3 205.1
E1 P1 4 172.8 E1 P2 4 168.5 E1 P3 4 189.5 E1 P4 4 186.1
E2 P5 1 202.8 E2 P6 1 196.8 E2 P7 1 209.6 E2 P8 1 195.6
E2 P5 2 200.2 E2 P6 2 205.3 E2 P7 2 209.1 E2 P8 2 207.0
E2 P5 3 207.9 E2 P6 3 206.6 E2 P7 3 203.2 E2 P8 3 203.6
E2 P5 4 190.1 E2 P6 4 194.7 E2 P7 4 197.3 E2 P8 4 204.9
E3 P9 1 187.2 E3 P10 1 203.7 E3 P11 1 201.2 E3 P12 1 190.6
E3 P9 2 197.8 E3 P10 2 193.6 E3 P11 2 196.5 E3 P12 2 196.1
E3 P9 3 191.9 E3 P10 3 194.4 E3 P11 3 197.8 E3 P12 3 202.9
E3 P9 4 199.5 E3 P10 4 197.8 E3 P11 4 191.5 E3 P12 4 202.5
E4 P13 1 193.6 E4 P14 1 206.3 E4 P15 1 186.0 E4 P16 1 172.2
E4 P13 2 204.2 E4 P14 2 202.0 E4 P15 2 172.2 E4 P16 2 194.0
E4 P13 3 206.3 E4 P14 3 176.1 E4 P15 3 187.5 E4 P16 3 187.8
E4 P13 4 202.0 E4 P14 4 173.5 E4 P15 4 179.0 E4 P16 4 200.1
;
ODS HTML FILE ="HIER";
ODS GRAPHICS ON;
PROC GLM DATA=HIER;
CLASS Especie Procedencia;
MODEL Producao= Especie Procedencia(Especie);
Test h=Especie e=Procedencia(Especie);
means Especie/TuKey e=Procedencia(Especie);
lsmeans Procedencia(Especie)/slice=Especie;
run;
ODS HTML CLOSE;
RUN;
QUIT;
```

Prog21. Experimento fracionário 2^{4-1}_{IV} . Exercício 6.5.4.1.

```

DATA FATFRAC;
INPUT TRAT X1 X2 X3 X4 PROD@@;
X12=X1*X2;
X13=X1*X3;
X23=X2*X3;
CARDS;
1 -1 -1 -1 43.0 2 1 -1 -1 1 95.6 3 -1 1 -1 1 43.0 4 1 1 -1 -1 62.1
5 -1 -1 1 1 71.8 6 1 -1 1 -1 57.4 7 -1 1 1 -1 76.5 8 1 1 1 1 91.8
1 -1 -1 -1 44.0 2 1 -1 -1 1 98.5 3 -1 1 -1 1 48.7 4 1 1 -1 -1 67.3
5 -1 -1 1 1 75.4 6 1 -1 1 -1 50.6 7 -1 1 1 -1 68.5 8 1 1 1 1 88.7
;
ODS HTML FILE ="FATFRAC";
ODS GRAPHICS ON;
PROC REG DATA=FATFRAC;
MODEL PROD= X1 X2 X3 X4 X12 X13 X23;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog22. Análise de grupos de experimentos. Exercício 7.3.1.

```

data GRUPO;
input Blocos Local$ Esp $ Prod@@;
cards;
1 L1 E1 35.7 1 L1 E2 29.9 1 L1 E3 42.2 1 L1 E4 44.9
2 L1 E1 20.6 2 L1 E2 33.1 2 L1 E3 48.9 2 L1 E4 50.4
3 L1 E1 33.0 3 L1 E2 33.9 3 L1 E3 38.4 3 L1 E4 45.0
4 L1 E1 33.4 4 L1 E2 46.2 4 L1 E3 28.4 4 L1 E4 54.6
5 L1 E1 20.9 5 L1 E2 27.2 5 L1 E3 37.8 5 L1 E4 55.1
1 L2 E1 40.4 1 L2 E2 29.3 1 L2 E3 42.7 1 L2 E4 67.2
2 L2 E1 34.9 2 L2 E2 41.3 2 L2 E3 23.9 2 L2 E4 65.3
3 L2 E1 35.0 3 L2 E2 25.4 3 L2 E3 36.8 3 L2 E4 53.0
4 L2 E1 44.6 4 L2 E2 45.1 4 L2 E3 33.7 4 L2 E4 69.4
5 L2 E1 45.1 5 L2 E2 29.8 5 L2 E3 47.8 5 L2 E4 59.4
1 L3 E1 40.2 1 L3 E2 26.1 1 L3 E3 27.0 1 L3 E4 51.3
2 L3 E1 42.5 2 L3 E2 25.5 2 L3 E3 21.8 2 L3 E4 41.5
3 L3 E1 49.3 3 L3 E2 33.7 3 L3 E3 41.2 3 L3 E4 55.7
4 L3 E1 39.9 4 L3 E2 13.1 4 L3 E3 31.3 4 L3 E4 64.5
5 L3 E1 49.1 5 L3 E2 26.0 5 L3 E3 30.4 5 L3 E4 56.7
1 L4 E1 28.3 1 L4 E2 31.2 1 L4 E3 42.1 1 L4 E4 48.6
2 L4 E1 23.6 2 L4 E2 41.8 2 L4 E3 47.5 2 L4 E4 48.7
3 L4 E1 27.9 3 L4 E2 27.6 3 L4 E3 38.2 3 L4 E4 51.0
4 L4 E1 34.7 4 L4 E2 32.8 4 L4 E3 38.9 4 L4 E4 38.8
5 L4 E1 13.6 5 L4 E2 27.9 5 L4 E3 45.5 5 L4 E4 49.1
;
ods html file="GRUPO";
ods graphics on;
proc glm data=GRUPO;
class Blocos Esp;
model Prod= Blocos Esp;
BY Local;
run;
proc glm;
class Blocos Local Esp;
model Prod= local Blocos(local) Esp Local*Esp/ss3;
Test H=Esp E=Local*Esp;
Test H=Local E=Local*Esp;
lsmeans local*Esp/slice=Esp;
lsmeans local*Esp/slice=Local;
run;
ods html close;
run;
quit;

```

Prog23. Análise de grupos de experimentos. Exercício 7.3.2.

```

data GRUPOBARBIN;
input Local $ Prog Blocos Prod@@;
cards;
L1 1 1 20.3 L1 1 2 19.6 L1 1 3 23.5 L1 1 4 19.1

```

```

L1 2 1 21.7 L1 2 2 19.3 L1 2 3 16.7 L1 2 4 18.5
L1 3 1 22.0 L1 3 2 24.9 L1 3 3 24.4 L1 3 4 20.8
L1 4 1 20.8 L1 4 2 23.0 L1 4 3 21.3 L1 4 4 24.9
L1 5 1 21.5 L1 5 2 22.3 L1 5 3 22.1 L1 5 4 21.9
L1 6 1 19.6 L1 6 2 17.7 L1 6 3 18.7 L1 6 4 22.0
L2 1 1 10.2 L2 1 2 11.7 L2 1 3 9.1 L2 1 4 8.1
L2 2 1 16.1 L2 2 2 10.8 L2 2 3 10.9 L2 2 4 10.3
L2 3 1 17.7 L2 3 2 13.1 L2 3 3 14.2 L2 3 4 11.0
L2 4 1 13.5 L2 4 2 14.4 L2 4 3 11.2 L2 4 4 12.8
L2 5 1 20.5 L2 5 2 12.5 L2 5 3 11.3 L2 5 4 12.2
L2 6 1 12.0 L2 6 2 13.0 L2 6 3 12.3 L2 6 4 10.6
L3 1 1 22.7 L3 1 2 21.4 L3 1 3 22.9 L3 1 4 22.0
L3 2 1 22.6 L3 2 2 21.4 L3 2 3 20.7 L3 2 4 20.8
L3 3 1 21.4 L3 3 2 21.7 L3 3 3 22.5 L3 3 4 19.4
L3 4 1 25.0 L3 4 2 23.6 L3 4 3 23.3 L3 4 4 24.8
L3 5 1 26.4 L3 5 2 26.4 L3 5 3 28.0 L3 5 4 27.3
L3 6 1 20.6 L3 6 2 23.5 L3 6 3 19.4 L3 6 4 21.9
;
ods html file="GRUPOBARBIN";
ods graphics on;
proc glm data=GRUPOBARBIN;
class Blocos Prog;
model Prod= Blocos Prog;
BY Local;
run;
proc glm;
class Blocos Local Prog;
model Prod= Blocos(local) local Prog Local*Prog/ss3;
Test H=Prog E=Local*Prog;
Test H=Local E=Local*prog;
lsmeans local*Prog/slice=Prog;
lsmeans local*Prog/slice=Local;
run;
ods html close;
run;
quit;

```

Prog24. Análise de experimento em parcelas subdivididas. Exercício 8.1.1.

```

data PSD;
input M $ A $ Blocos Prod @@;
cards;
M1 A1 1 58.1 M1 A1 2 58.4 M1 A1 3 39.5 M1 A1 4 44.6
M1 A2 1 72.7 M1 A2 2 77.9 M1 A2 3 58.5 M1 A2 4 61.7
M1 A3 1 65.9 M1 A3 2 71.7 M1 A3 3 54.2 M1 A3 4 55.5
M1 A4 1 59.9 M1 A4 2 55.7 M1 A4 3 37.7 M1 A4 4 46.2
M2 A1 1 71.5 M2 A1 2 92.7 M2 A1 3 60.5 M2 A1 4 46.7
M2 A2 1 76.7 M2 A2 2 93.7 M2 A2 3 56.5 M2 A2 4 69.1
M2 A3 1 79.7 M2 A3 2 87.7 M2 A3 3 55.1 M2 A3 4 60.5
M2 A4 1 85.4 M2 A4 2 76.5 M2 A4 3 58.7 M2 A4 4 68.7
M3 A1 1 83.0 M3 A1 2 77.9 M3 A1 3 59.4 M3 A1 4 67.0
M3 A2 1 84.5 M3 A2 2 67.1 M3 A2 3 59.9 M3 A2 4 62.2
M3 A3 1 85.9 M3 A3 2 64.4 M3 A3 3 83.4 M3 A3 4 67.0
M3 A4 1 85.7 M3 A4 2 74.7 M3 A4 3 72.2 M3 A4 4 69.0
M4 A1 1 99.5 M4 A1 2 87.4 M4 A1 3 71.9 M4 A1 4 70.2
M4 A2 1 93.7 M4 A2 2 89.7 M4 A2 3 76.7 M4 A2 4 77.9
M4 A3 1 91.0 M4 A3 2 87.0 M4 A3 3 63.7 M4 A3 4 67.9
M4 A4 1 95.4 M4 A4 2 92.0 M4 A4 3 75.4 M4 A4 4 63.1
;
ods html file="PSD";
ods graphics on;
proc GLM;
class Blocos M A;
model Prod=Blocos M Blocos*M A M*A;
TEST H=Blocos M E=blocos*M;
lsmeans M*A/slice=M;
lsmeans M*A/slice=A;
run;
ods html close;
run;
quit;

```

Prog25. Análise de experimento em faixas. Exercício 8.2.1.

```

data FAIXA;
input A $ B $ Blocos Prod @@;
cards;
A1 B1 1 152.4 A1 B2 1 190.8 A1 B3 1 173.1 A1 B4 1 157.2
A1 B1 2 153.3 A1 B2 2 204.6 A1 B3 2 188.4 A1 B4 2 146.1
A1 B1 3 103.5 A1 B2 3 153.6 A1 B3 3 142.2 A1 B4 3 99.0
A1 B1 4 117.0 A1 B2 4 162.0 A1 B3 4 145.8 A1 B4 4 121.2
A2 B1 1 187.8 A2 B2 1 201.3 A2 B3 1 209.4 A2 B4 1 224.4
A2 B1 2 243.3 A2 B2 2 246.0 A2 B3 2 230.4 A2 B4 2 201.0
A2 B1 3 158.7 A2 B2 3 148.2 A2 B3 3 144.6 A2 B4 3 154.2
A2 B1 4 122.7 A2 B2 4 181.5 A2 B3 4 158.7 A2 B4 4 180.3
A3 B1 1 218.1 A3 B2 1 222.0 A3 B3 1 225.6 A3 B4 1 225.0
A3 B1 2 204.6 A3 B2 2 176.1 A3 B3 2 169.2 A3 B4 2 196.2
A3 B1 3 156.0 A3 B2 3 157.2 A3 B3 3 219.0 A3 B4 3 189.6
A3 B1 4 175.8 A3 B2 4 163.2 A3 B3 4 175.8 A3 B4 4 181.2
A4 B1 1 261.3 A4 B2 1 246.0 A4 B3 1 239.1 A4 B4 1 250.5
A4 B1 2 229.5 A4 B2 2 235.5 A4 B3 2 228.6 A4 B4 2 241.5
A4 B1 3 188.7 A4 B2 3 201.3 A4 B3 3 167.1 A4 B4 3 198.0
A4 B1 4 184.2 A4 B2 4 204.6 A4 B3 4 178.2 A4 B4 4 165.6
;
ods html file="FAIXA";
ods graphics on;
proc GLM DATA= FAIXA;
class Blocos A B;
model Prod=blocos A Blocos*A B Blocos*B A*B;
TEST H=Blocos A E=Blocos*A;
TEST H=B E=Blocos*B;
run;
ods html close;
run;
quit;

```

Prog26. Análise BIB. Tipo I. Análise intrablocos. Exercício 9.3.1.1.

```

DATA BIBTI;
INPUT REP BLOCO TRAT $ PROD@@;
CARDS;
1 1 M1 36.6 1 1 M2 32.9 1 2 M3 27.4
1 2 M4 29.7 1 3 M5 25.6 1 3 M6 27.4
2 4 M1 43.9 2 4 M3 32.8 2 5 M2 45.7
2 5 M5 29.2 2 6 M4 23.7 2 6 M6 21.9
3 7 M1 42.0 3 7 M4 31.1 3 8 M2 47.5
3 8 M6 31.7 3 9 M3 27.4 3 9 M5 23.7
4 10 M1 38.4 4 10 M5 23.7 4 11 M2 42.0
4 11 M4 18.3 4 12 M3 29.4 4 12 M6 23.7
5 13 M1 51.2 5 13 M6 25.6 5 14 M2 49.4
5 14 M3 33.9 5 15 M4 32.5 5 15 M5 29.5
;
ods html file="BIBTI";
ods graphics on;
PROC PRINT;
PROC GLM DATA=BIBINTRATIPOI;
CLASS REP BLOCO TRAT;
MODEL PROD=REP BLOCO(REP) TRAT/SS1;
RUN;
LSMEANS TRAT/PDIFF=ALL ADJUST=TUKEY;
RUN;
ods html close;
run;
quit;

```

Prog27. Análise BIB. Tipo II. Análise intrablocos. Exercício 9.3.1.2

```

DATA BIBTII;
INPUT GRUPO $ BLOCO TRAT $ PROD@@;
CARDS;
A 1 E1 81.5 A 2 E2 74.5 A 4 E4 65.2
A 3 E3 81.5 A 5 E5 62.9 A 6 E6 55.9
A 7 E7 72.2 A 1 E2 65.2 A 2 E3 86.2
A 3 E4 58.2 A 4 E5 55.9 A 5 E6 74.5
A 6 E7 60.5 A 7 E1 62.9 B 8 E1 88.5
B 9 E3 83.8 B 10 E5 46.6 B 11 E7 65.2

```

```

B 12 E2 60.5 B 13 E4 53.5 B 14 E6 65.2
B 8 E3 93.2 B 9 E5 55.9 B 10 E7 69.9
B 11 E2 58.2 B 12 E4 65.2 B 13 E6 55.9
B 14 E1 76.8 C 15 E1 69.9 C 16 E4 62.9
C 17 E7 74.5 C 18 E3 76.8 C 19 E6 65.2
C 20 E2 67.5 C 21 E5 46.6 C 15 E4 51.2
C 16 E7 79.2 C 17 E3 90.8 C 18 E6 55.9
C 19 E2 79.2 C 20 E5 53.5 C 21 E1 76.8
;
ods html file="BIBTII";
ods graphics on;
PROC GLM;
CLASS GRUPO BLOCO TRAT;
MODEL PROD = GRUPO BLOCO(GRUPO) TRAT/SS1;
RUN;
LSMEANS TRAT/PDIFF;
RUN;
ods html CLOSE;
RUN;
QUIT;

```

Prog28. Análise BIB. Tipo III. Análise intrablocos. Exercício 9.3.1.3

```

DATA BIBTIII;
INPUT BLOCO TRAT $ PROD@@;
CARDS;
1 P3 54.0 2 P3 49.1 3 P10 34.0 4 P2 58.3
5 P7 49.9 6 P4 65.3 7 P1 74.1 8 P3 73.4
9 P1 81.5 10 P2 61.2 11 P1 78.1 12 P1 67.9
13 P2 64.7 1 P6 42.4 2 P4 42.2 3 P11 31.2
4 P5 57.6 5 P8 70.1 6 P5 69.1 7 P5 66.4
8 P5 69.1 9 P2 70.2 10 P4 65.5 11 P4 66.4
12 P6 71.9 13 P6 67.2 1 P9 61.9 2 P8 76.1
3 P12 67.6 4 P8 81.0 5 P9 65.7 6 P6 58.0
7 P9 54.8 8 P7 71.1 9 P3 79.6 10 P9 57.4
11 P7 66.4 12 P8 64.3 13 P7 83.9 1 P11 42.5
2 P12 48.4 3 P13 66.8 4 P11 27.1 5 P10 69.1
6 P10 70.0 7 P12 65.1 8 P13 88.8 9 P10 66.8
10 P13 85.3 11 P11 50.6 12 P13 97.7 13 P12 57.0
;
ods html file="BIBTIII";
ods graphics on;
PROC GLM DATA=BIBTIII;
CLASS BLOCO TRAT;
MODEL PROD= BLOCO TRAT/SS1;
RUN;
LSMEANS TRAT/PDIFF=ALL ADJUST=TUKEY;
RUN;
ods html close;
run;
quit;

```

Prog29. Látice quadrado (Square Lattice). Análise intrablocos. Exercício 10.2.1.1.

```

DATA LATICEINTRA1;
INPUT REP BLOCO TRAT $ PROD@@;
CARDS;
1 1 E1 79.6 1 1 E2 66.6 1 1 E3 84.9 1 2 E4 74.2
1 2 E5 30.7 1 2 E6 67.3 1 3 E7 46.0 1 3 E8 65.8
1 3 E9 63.7 2 4 E1 79.6 2 4 E4 55.7 2 4 E7 49.0
2 5 E2 68.2 2 5 E5 58.9 2 5 E8 65.8 2 6 E3 69.0
2 6 E6 61.3 2 6 E9 49.2 3 7 E1 64.5 3 7 E5 35.0
3 7 E9 33.5 3 8 E2 53.1 3 8 E6 65.2 3 8 E7 54.3
3 9 E3 56.6 3 9 E4 47.4 3 9 E8 34.1 4 10 E1 67.3
4 10 E6 59.1 4 10 E8 50.5 4 11 E2 48.2 4 11 E4 56.1
4 11 E9 60.0 4 12 E3 57.4 4 12 E5 39.2 4 12 E7 31.4
;
ods html file="LATICEINTRA1";
ods graphics on;
PROC GLM DATA=LATICEINTRA1;
CLASS REP BLOCO TRAT;
MODEL PROD=REP BLOCO(REP) TRAT;
LSMEANS TRAT/PDIFF=ALL ADJUST=TUKEY;
RUN; ODS HTML CLOSE;RUN;QUIT;

```

Prog30. Látice quadrado (Square Lattice). Análise intrablocos. Exercício 10.2.1.2.

```

DATA LATICEINTRA2;
INPUT REP BLOCO TRAT $ PROD @@;
CARDS;
1 1 E1 31.9 1 2 E5 42.5 1 3 E9 24.6 1 4 E13 34.8
1 1 E2 38.5 1 2 E6 38.2 1 3 E10 48.8 1 4 E14 51.4
1 1 E3 35.4 1 2 E7 21.1 1 3 E11 23.7 1 4 E15 37.3
1 1 E4 59.0 1 2 E8 27.6 1 3 E12 32.0 1 4 E16 42.9
2 5 E1 40.2 2 6 E2 53.7 2 7 E3 47.3 2 8 E4 60.5
2 5 E5 42.0 2 6 E6 43.4 2 7 E7 31.3 2 8 E8 42.0
2 5 E9 40.8 2 6 E10 39.7 2 7 E11 33.7 2 8 E12 31.7
2 5 E13 23.9 2 6 E14 31.0 2 7 E15 36.9 2 8 E16 23.6
3 9 E1 45.3 3 10 E2 28.2 3 11 E3 25.7 3 12 E4 66.5
3 9 E6 43.8 3 10 E5 28.7 3 11 E8 40.4 3 12 E7 55.9
3 9 E11 39.3 3 10 E12 37.0 3 11 E9 24.2 3 12 E10 49.9
3 9 E16 46.9 3 10 E15 37.8 3 11 E14 36.3 3 12 E13 33.4
;
ods html file="LATICEINTRA2";
ods graphics on;
PROC GLM DATA=LATICEINTRA2;
CLASS REP BLOCO TRAT;
MODEL PROD=REP BLOCO(REP) TRAT;
LSMEANS TRAT/PDIFF=ALL ADJUST=TUKEY;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog31. Látice quadrado (Square Lattice). Análise com recuperação da informação interblocos. Exercício 10.2.2.1.

```

DATA LATICEINTER1;
INPUT GROUP BLOCK TREATMNT PROD @@;
CARDS;
1 1 1 79.6 1 2 4 74.2 1 3 7 46.0 1 1 2 66.6
1 2 5 30.7 1 3 8 65.8 1 1 3 84.9 1 2 6 67.3
1 3 9 63.7 2 1 1 79.6 2 2 2 68.2 2 3 3 69.0
2 1 4 55.7 2 2 5 58.9 2 3 6 61.3 2 1 7 49.0
2 2 8 65.8 2 3 9 49.2 3 1 1 64.5 3 2 2 53.1
3 3 3 56.6 3 1 5 35.0 3 2 6 65.2 3 3 4 47.4
3 1 9 33.5 3 2 7 54.3 3 3 8 34.1 4 1 1 67.3
4 2 2 48.2 4 3 3 57.4 4 1 6 59.1 4 2 4 56.1
4 3 5 39.2 4 1 8 50.5 4 2 9 60.0 4 3 7 31.4
;
ods html file="LATICEINTER1";
ods graphics on;
PROC LATTICE DATA=LATICEINTER1;
VAR PROD;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog32. Látice quadrado (Square Lattice). Análise com recuperação da informação interblocos. Exercício 10.2.2.2.

```

DATA LATICEINTER2;
INPUT GROUP BLOCK TREATMNT PROD @@;
CARDS;
1 1 1 31.9 1 1 2 38.5 1 1 3 35.4 1 1 4 59.0
1 2 5 42.5 1 2 6 38.2 1 2 7 21.1 1 2 8 27.6
1 3 9 24.6 1 3 10 48.8 1 3 11 23.7 1 3 12 32.0
1 4 13 34.8 1 4 14 51.4 1 4 15 37.3 1 4 16 42.9
2 1 1 40.2 2 1 5 42.0 2 1 9 40.8 2 1 13 23.9
2 2 2 53.7 2 2 6 43.4 2 2 10 39.7 2 2 14 31.0
2 3 3 47.3 2 3 7 31.3 2 3 11 33.7 2 3 15 36.9
2 4 4 60.5 2 4 8 42.0 2 4 12 31.7 2 4 16 23.6
3 1 1 45.3 3 1 6 43.8 3 1 11 39.3 3 1 16 46.9
3 2 2 28.2 3 2 5 28.7 3 2 12 37.0 3 2 15 37.8
3 3 3 25.7 3 3 8 40.4 3 3 9 24.2 3 3 14 36.3
3 4 4 66.5 3 4 7 55.9 3 4 10 49.9 3 4 13 33.4
;

```

```
ods html file="LATICEINTER2";
ods graphics on;
PROC LATTICE DATA=LATICEINTER2;
VAR PROD;
RUN;
ODS HTML CLOSE;
RUN; QUIT;
```

Prog33. Análise de regressão fatorial 3×3

```
DATA FATORIAL33;
INPUT TRAT $ REP P K PROD@@;
CARDS;
P0K0 1 -1 -1 45.6 P1K2 1 0 1 52.5
P2K1 1 1 0 69.2 P0K1 1 -1 0 84.2
P1K0 1 0 -1 71.0 P2K2 1 1 1 83.8
P0K2 1 -1 1 34.9 P1K1 1 0 0 102.2
P2K0 1 1 -1 62.9 P0K0 2 -1 -1 49.6
P1K2 2 0 1 62.7 P2K1 2 1 0 103.5
P0K1 2 -1 0 80.0 P1K0 2 0 -1 33.0
P2K2 2 1 1 61.0 P0K2 2 -1 1 41.1
P1K1 2 0 0 103.2 P2K0 2 1 -1 59.2
;
ODS HTML FILE ="FATORIAL33";
ODS GRAPHICS ON;
PROC RSREG DATA=FATORIAL33;
MODEL PROD=P K;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;
```

Prog34. Operações com matrizes. Fatorial 3×3. Exercício 11.1.5.1.

```
proc IML;
reset print;
M={-12.366667 5.2375, 5.2375 -71.216666};
H={-6.183333 2.61875, 2.61875 -35.608333};
C={-8.683333, -1.225};
CE={-2.683333, 4.775};
INVM=INV(M);
COFE=INVM*CE;
COF=INVM*C;
AVALA=EIGVAL(H);
AVECA=EIGVEC(H);
A={-6.183333 2.61875, 2.61875 -35.608333};
P={0.9961239 -0.0879612, 0.0879612 0.9961239};
D={-5.952089 0, 0 -35.839580};
PT=TRANSP(P);
AA=P*D*PT;
DD=PT*A*P;
QUIT;
```

Prog35. Superfície de resposta fatorial 3×3×3. Exercício 11.2.3.1

```
DATA FATORIAL333;
INPUT TRAT $ REP N P K PROD@@;
CARDS;
N0P0K0 1 -1 -1 -1 45.6 N0P1K2 1 -1 0 1 52.5 N0P2K1 1 -1 1 0 69.2
N1P0K1 1 0 -1 0 57.0 N1P1K0 1 0 0 -1 66.3 N1P2K2 1 0 1 1 71.3
N2P0K2 1 1 -1 1 44.6 N2P1K1 1 1 0 0 43.3 N2P2K0 1 1 1 -1 60.4
N0P0K1 1 -1 -1 0 84.2 N0P1K0 1 -1 0 -1 71.0 N0P2P2 1 -1 1 1 83.8
N1P0K2 1 0 -1 1 43.3 N1P1K1 1 0 0 0 98.2 N1P2K0 1 0 1 -1 69.2
N2P0K0 1 1 -1 -1 69.8 N2P1K2 1 1 0 1 49.3 N2P2K1 1 1 1 0 83.2
N0P0K2 1 -1 -1 1 34.9 N0P1K1 1 -1 0 0 92.2 N0P2K0 1 -1 1 -1 62.9
N1P0K0 1 0 -1 -1 35.9 N1P1K2 1 0 0 1 88.2 N1P2K1 1 0 1 0 107.7
N2P0K1 1 1 -1 0 60.9 N2P1K0 1 1 0 -1 40.7 N2P2K2 1 1 1 1 55.5
N0P0K0 2 -1 -1 -1 36.1 N0P1K2 2 -1 0 1 45.6 N0P2K1 2 -1 1 0 75.3
N1P0K1 2 0 -1 0 42.2 N1P1K0 2 0 0 -1 85.6 N1P2K2 2 0 1 1 75.1
N2P0K2 2 1 -1 1 43.4 N2P1K1 2 1 0 0 65.5 N2P2K0 2 1 1 -1 56.7
N0P0K1 2 -1 -1 0 58.2 N0P1K0 2 -1 0 -1 24.0 N0P2P2 2 -1 1 1 44.4
N1P0K2 2 0 -1 1 29.7 N1P1K1 2 0 0 0 97.7 N1P2K0 2 0 1 -1 75.8
N2P0K0 2 1 -1 -1 37.6 N2P1K2 2 1 0 1 46.8 N2P2K1 2 1 1 0 50.7
```

```

N0P0K2 2 -1 -1 1 29.9 N0P1K1 2 -1 0 0 61.5 N0P2K0 2 -1 1 -1 43.1
N1P0K0 2 0 -1 -1 32.2 N1P1K2 2 0 0 1 72.6 N1P2K1 2 0 1 0 90.0
N2P0K1 2 1 -1 0 39.6 N2P1K0 2 1 0 -1 82.9 N2P2K2 2 1 1 1 60.4
;
ODS HTML FILE ="FATORIAL333";
ODS GRAPHICS ON;
PROC RSREG DATA=FATORIAL333;
MODEL PROD=N P K;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog36. Superfície de resposta. Delineamento composto central. Exercício 11.3.1.1.

```

DATA DCC;
INPUT TRAT $ REP N P K PROD@@;
CARDS;
111 1 -1 -1 19.1 113 1 -1 -1 1 15.4 131 1 -1 1 -1 24.0
133 1 -1 1 1 32.7 311 1 1 -1 -1 29.0 313 1 1 -1 1 25.3
331 1 1 1 -1 27.7 333 1 1 1 1 32.0 222 1 0 0 0 22.8
422 1 2 0 0 12.3 022 1 -2 0 0 20.9 242 1 0 2 0 12.9
202 1 0 -2 0 14.1 224 1 0 0 2 17.8 220 1 0 0 -2 18.5
111 2 -1 -1 -1 25.9 113 2 -1 -1 1 18.5 131 2 -1 1 -1 29.6
133 2 -1 1 1 21.3 311 2 1 -1 -1 18.5 313 2 1 -1 1 29.6
331 2 1 1 -1 29.6 333 2 1 1 1 29.6 222 2 0 0 0 30.8
422 2 2 0 0 18.5 022 2 -2 0 0 12.9 242 2 0 2 0 16.0
202 2 0 -2 0 12.3 224 2 0 0 2 16.0 220 2 0 0 -2 17.8
;
ODS HTML FILE ="DCC";
ODS GRAPHICS ON;
PROC RSREG DATA=DCC;
MODEL PROD=N P K;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;

```

Prog37. Operações com matrizes. Delineamento composto central. Exercício 11.3.1.1

```

proc IML;
reset print;
M={-6.8612 -0.7625 1.4, -0.7625 -8.0236 0.525, 1.4 0.525 -6.1736};
H={-3.4306 -0.38125 0.7, -0.38125 -4.0118 0.2625, 0.7 0.2625 -3.0868};
C={-0.9, -1.5688, 0.125};
CE={5.1, 4.4312, 6.125};
TITLE 'Inversa de M';
INVM=INV(M);
COFE=INVM*C;
COF=INVM*C;
title 'Autovalores de H';
AVALH=EIGVAL(H);
itle 'Autovetores de H';
AVECA=EIGVEC(H);
A={-3.4306 -0.38125 0.7, -0.38125 -4.0118 0.2625, 0.7 0.2625 -3.0868};
P={0.6223459 -0.553866 0.5530988, -0.021619 0.6941842 0.7194725, 0.7824437 0.4597184 -0.420049};
D={-2.537282, -3.533775, -4.458143};
PT=TRANSP(P);
AA=P*D*PT;
DD=PT*A*P;
PTP=P*PT;
QUIT;

```

Prog38. Delineamento Box-Behnken. Exercício 11.4.1.

```

DATA BOXBEHNKEN;
INPUT TRAT N P K PROD@@;
CARDS;
1 0 0 60 47.3 2 120 0 60 45.2 3 0 120 60 58.1
4 120 120 60 57.0 5 0 60 0 50.7 6 120 60 0 52.5
7 0 60 120 57.3 8 120 60 120 52.2 9 60 0 0 45.8
10 60 120 0 52.1 11 60 0 120 42.2 12 60 120 120 58.8

```

```

13 60 60 60 55.1 14 60 60 60 56.3 15 60 60 60 57.1
16 60 60 60 61.8
;
ODS HTML FILE ="BOXBEHNKEN";
ODS GRAPHICS ON;
PROC RSREG DATA=BOXBEHNKEN;
MODEL PROD=N P K;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;
RUN;
proc IML;
reset print;
A={-1.1125 0.125 -0.8625, 0.125 -4.5625 1.2875, -0.8625 1.2875 -3.2875};
B={-0.8125, 5.6875, 1.175};
INVA=INV(A);
C=-0.5*(INVA)*B;
AVALA=EIGVAL(A);
AVECA=EIGVEC(A);
QUIT;

```

Prog39. Aplicação de um fatorial $(\frac{1}{5})5^3$. Exercício 11.5.1.

```

DATA FATFRAC;
INPUT N P K PROD@@@;
CARDS;
0 0 0 60.5 30 30 30 134.5 60 60 60 168.9 90 90 90 199.9
120 120 120 183.0 30 90 120 158.7 60 120 0 142.1 90 0 30 126.7
120 30 60 140.4 0 60 90 148.4 60 30 90 158.8 90 60 120 184.1
120 90 0 144.5 0 120 30 122.7 30 0 60 108.1 90 120 60 165.3
120 0 90 147.1 0 30 120 108.7 30 60 0 155.7 60 90 30 186.7
120 60 30 184.6 0 90 60 155.7 30 120 90 179.1 60 0 120 135.9
90 30 0 115.3
;
ODS HTML FILE ="FATFRAC";
ODS GRAPHICS ON;
PROC RSREG DATA=FATFRAC;
MODEL PROD=N P K;
RUN;
ODS HTML CLOSE;
RUN;
QUIT;
proc IML;
reset print;
A={-18.165176 -1.975132 4.708943, -1.975132 -29.499094 -0.621885, 4.708943 -0.621885 -14.781316};
B={18.496, 24.624, 15.308};
INVA=INV(A);
C=-0.5*(INVA)*B;
AVALA=EIGVAL(A);
AVECA=EIGVEC(A);
QUIT;

```