UNIVERSIDADE ESTADUAL DA PARAIBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA

JOÃO BATISTA FILGUEIRA COSTA

ASPECTOS TEÓRICOS E PRÁTICOS DA ANÁLISE ESTATÍSTICA DE UM EXPERIMENTO EM BLOCOS CASUALIZADOS COM EFEITOS ALEATÓRIOS

CAMPINA GRANDE
DEZEMBRO DE 2010

JOÃO BATISTA FILGUEIRA COSTA

ASPECTOS TEÓRICOS E PRÁTICOS DA ANÁLISE ESTATÍSTICA DE UM EXPERIMENTO EM BLOCOS CASUALIZADOS COM EFEITOS ALEATÓRIOS

Trabalho de Conclusão de Curso (TCC) apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientador:
JOÃO GIL DE LUNA

CAMPINA GRANDE
DEZEMBRO DE 2010

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL - UEPB

C837a Costa, João Batista Filgueira.

Aspectos teóricos e práticos da análise estatística de um experimento em blocos casualizados com efeitos aleatórios [manuscrito] / João Batista Filgueira Costa. - 2010.

46 f.: il.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologias, 2010.

”Orientação: Prof. Dr. João Gil de Luna, Departamento de Estatística”.

1. Estatística aplicada. 2. Análise estatística de dados. 3. Estatística. I. Título.

21. ed. CDD 519.53

JOÃO BATISTA FILGUEIRA COSTA

ASPECTOS TEÓRICOS E PRÁTICOS DA ANÁLISE ESTATÍSTICA DE UM EXPERIMENTO EM BLOCOS CASUALIZADOS COM EFEITOS ALEATÓRIOS

Trabalho Acadêmico Orientado apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Aprovado em: 13/12/2010

Banca Examinadora:

Nota 8,0

Prof. João Gil de Luna
Orientador
Universidade Estadual da Paraíba

Prof. Juarez Fernandes de Oliveira
Universidade Estadual da Paraíba

Prof. Fernando Bezerra Cavalcanti
Universidade Estadual da Paraíba

Dedicatória

A minha filha Guiomar, que fez-me abrir os olhos para o tempo, refletir e descobrir ainda que tardio, que é possível sonhar e realizar os mais impossíveis sonhos desde que haja dedicação e ética.

Agradecimentos

Agradeço a minha esposa(Girlene Filgueira), aos meus pais(Humberto Saturnino e Cecília Filgueira), meus irmãos e irmães, aos meus professores em particular (PHD. João Gil de Luna; Fernando Bezerra e Dr. Edwirde) e aos demais professores os quais tenho muito respeito admiração e estima, ao coordenador do curso de estatística Prof.(Dr. Gustavo Esteves) aos meus colegas de turma(Thiago, Emanuela, Marcelo Sampaio), aos funcionários(Vera na biblioteca, Nadilma na secretaria), ao diretor do Centro de Ciências Tecnológicas(CCT) Prof. Juarez Fernandes e Prof. José Tavares por colaborarem e subsidiarem no período em que estive à presidência do Centro Acadêmico de Estatística(CAE), minhas tias(Inalda Saturnino e Marinês Barros), ao casal(Hailton Saturnino e Zetinha dos Anjos) ao meu primo (Cláudio Germano).

Resumo

Geralmente os textos disponíveis na literatura especializada trata a *Estatística Experimental* como um conjunto de procedimentos que se propõe a resolver os problemas inferenciais das pesquisas científicas em outras áreas do conhecimento e pouco, ou quase nada, está interessada em aprofundar-se no formalismo matemático/estatístico que dão origem a estes procedimentos. Este trabalho se propõe a contribuir com o desenvolvimento teórico dos procedimentos que dão suporte às análises estatísticas dos dados de um experimento em blocos casualizados de efeitos aleatórios, uma vez que esse enfoque não é encontrado com frequência na literatura ou ocorre de forma bastante fragmentada. Neste sentido, procurou-se apresentar alguns conceitos utilizados na experimentação, definiu-se o modelo matemático associado às observações experimentais, encontrou-se os estimadores de máxima verossimilhança dos efeitos dos fatores sobre a variável resposta, discutiu-se algumas propriedades dos estimadores e os argumentos teóricos que fundamentam a análise da variância (ANOVA). Além disso, utilizou-se o método da ANOVA para estimar os componentes da variância discutindo-se as análises complementares. Os dados de um experimento real foram analisados e os resultados foram comentados detalhadamente.

Palavras-chave - Análise de Variância, Componente de Variância, Modelo Aleatório.

Abstract

The texts available in specific literature generally treats the experimental statistic like a set of the procedures that proposes to solve the inferences problems of the scientific researches in others areas of the knowledge and hardly ever to be interested in to study in depth in mathematic/statistic formalism that gives origin to theses procedures. This work proposes to contribute with the theoretical development of the procedures that gives support to the statistics analysis of the experiment in blocks by chance of the Random process, because this focus doesn't is found in frequency in the literature or to happen of the form enough fragmentary. Therefore it look for present some concept used in the experimentation, defined the mathematic model associated to the experimental observations, it found the estimators of the maximum verisimilitude of the factories' effect about the answer variable, it discussed some estimator´s property and the theoretical argument that found the analysis of variance (ANOVA). Besides which, it used the ANOVA method to estimate the complementary analyses. The factor of the one real experiment was analyzed and the results were commented on details.

Key words - Analysis of Variance, Variance Components, Random model.

Sumário

Lista de Tabelas

Lista de abreviaturas

1 Introdução p. 10

1.1 Princípios básicos da experimentação . . . p. 12

1.1.1 Repetição . . . p. 12

1.1.2 Casualização . . . p. 12

1.1.3 Controle local . . . p. 12

1.2 Conceitos do plano experimental . . . p. 14

1.2.1 Experimento . . . p. 14

1.2.2 Tratamento . . . p. 14

1.2.3 Testemunha . . . p. 15

1.2.4 Resíduo experimental . . . p. 15

1.2.5 Parcela . . . p. 15

1.2.6 Considerações Gerais do Delineamento em Blocos Casualizados . . . p. 16

1.2.6.1 Vantagens na instalação de um Delineamento em Blocos Casualizados . . . p. 16

1.2.6.2 Desvantagens na instalação de um Delineamento em Blocos Casualizados . . . p. 17

1.3 Componentes de variância no modelo aleatório . . . p. 17

1.3.1 Definição . . . p. 17

2 Desenvolvimento teórico p. 18

2.1 Efeitos fixos, efeitos aleatórios e hipóteses de interesse . . . . p. 18

2.2 O modelo estatístico . . . . p. 19

2.3 Estimação dos efeitos do modelo . . . . p. 20

2.4 Decomposição da variabilidade total . . . . p. 23

2.5 Distribuições de probabilidade dos estimadores dos efeitos dos componentes do modelo e das somas dos quadrados . . . . p. 23

2.6 Valores esperados das somas dos quadrados e dos quadrados médios . . . . . p. 29

2.7 Análises estatísticas . . . . p. 32

2.8 Estimação por ponto dos componentes da variância . . . . p. 34

2.8.1 Estimadores da proporção da variabilidade total explicada pelos componente da variância . . . . p. 35

2.8.2 Estimativa negativa de um componente da variância . . . . p. 35

2.9 Estimação por intervalo dos componentes da variância . . . . p. 35

2.10 Procedimento prático para a análise dos dados de um experimento em blocos ao acaso de efeitos aleatórios . . . . p. 36

2.11 Uma aplicação da teoria a uma pesquisa real . . . . p. 38

2.11.1 Cálculos das somas dos quadrados e a análise da variância . . . . . . p. 38

2.11.2 Estimativas dos componentes da variância . . . . p. 39

2.11.3 Estimativas das proporções da variabilidade total explicada pelos componentes da variância . . . . p. 40

2.11.4 Estimativas dos intervalos com 95% de confiança para os componentes da variância . . . . p. 40

3 Conclusões p. 42

Referências p. 44

Lista de Tabelas

1 Entrada de dados de um experimento em blocos completos casualizados. . . . p. 14

2 Distribuições de probabilidades dos estimadores dos efeitos do modelo e das somas dos quadrados . . . . p. 29

3 Tabela da análise de variância com as esperaças dos quadrados médios . . . . p. 32

4 Dados da produão de grãos (*ton*/*ha*) de 10 cultivares de arroz obtidos de um experimento em blocos casualizados com tês repetições, Lavras-MG. . . . . . p. 38

5 A análise da variância . . . . p. 39

Lista de abreviaturas

ANOVA (ANalysis Of VAriance)

FV - Fonte de Variação, ou seja, as partes da variação total

GL - Graus de liberdade associados à FV

SQ - Soma de quadrados (variação devida FV)

QM - Quadrado médio (Quociente da SQ pelo GL)

F - Valor da estatística de teste F de Fisher

ML - (máxima verossimilhança)

1 Introdução

Um dos primeiros registros sobre a realzação de experimentos agrícolas deu-se há mais de três séculos, pelo Belga Jonhan Baptista Vam Helmont. O objeto de estudo de Helmont baseava-se no fato de mostrar que os vegetais eram formados somente por água. Despresando outros componentes constituídos pelo solo ou ambiemte onde estes fossem cultivados. Helmont foi o pioneiro pesquisador a empregar a técnica de **causas controladas,** em que se observou com registro, analisando três variáveis que foram: uma planta, uma unidade de área limitada, ou parcela e uma ilimitada quantidade de água ou tratamento, no período de cinco anos. Atualmente um estudo dessa natureza onde se realiza uma simples observação por causas controladas, de modo sistemático, chama-se experimento. Em 1834, surge na França a primeira estação experimental, Boussingault, fundada por Busenka, com a finalidade específica de estudar as fontes de fornecimento de nitrogênio para as plantas e outros nutrientes nos manejos rotacionais, desta estação experimental surge investigações consideradas obras primas de uma pesquisa experimental. Santos(2008, p.23).

Em meados do século XIX houve um grande avanço na área das ciências agronômicas, em virtude da necessidade para implantação de técnicas e métodos que viabilizasse a maximização da produção agrícola em larga escala, assim como novas variedades de plantas com teor de produção elevado obdecendo as pespectivas futuras sobre alimentação.

Para atender as necessidades da produção agrícola foram criadas as estações experimentais, que no decorrer dos anos seguintes foi a base para todo desenvolvimento e aplicação dos termos e técnicas expecíficos, chamados de experimentos . No inicio do século XX com o desenvolvimento da indústria petrolífera surge da química substâncias necessárias e suficiente para a fertilização dos solos, aplicando nutrientes ausentes com o objetivo de equilibrar o potencial produtivo, necessitando de estudos sistematizados dos experimentos, ou seja, o planejamento.

Com os estudos de Ronald Fisher chefe da estação experimental **Rothamsted**, na Inglaterra no périodo de 1912 a 1933. Suas idéias revolucionaram o campo da experimentação, eliminando o carater empírico, ou seja, técnica conhecida apenas na observação, desprezando no entanto o estudo. Fisher desenvolveu uma técnica conhecida na estatística como análise de variância. Esta técnica teve enorme repercursão entre os cientistas e causou um passo extraodinário não somente no planejamento de experimentos mas também na análise e interpretação dos resultados, cujas técnicas são utilizadas até os dias atuais.

Os pesquisadores dispõem de uma gama de Softweres Estatisticos, uma das ferramentas usadas para análise de dados, que diante do domínio devido, efetuará cálculos complexos em fração de segundos. Existe atualmente um grande volune de trabalhos publicados na área da pesquisa agronômica, mas estar sendo notável a aplicação sutil em todos os ramos do conhecimento humano, e esse avanço é decorrente em parte, ás técnicas e teorias da estatística para validação cientifica desses estudos. Quando um pesquisador realiza um estudo, no decorrer do tempo se depara com um problema peculiar generalizado, que é a ausência de fontes bibliográficas que subsidie o desenvolvimento estatístico teórico exploratório. Quando em um experimento estamos interessados em estudar apenas um tipo de variável independente dizemos que possuimos apenas um fator. O tipo de tratamento tem importância significativa na forma como os dados serão analisados, influênciando nas devidas conclusões posteriores. Os tratamentos também são chamados de variáveis independentes.

As conclusões que se obtém na analise de experimentos com efeitos de tratamentos fixos são diferentemente das conclusões que se obtém na análise de experimentos com efeitos de tratamentos aleatórios. Se os efeitos são fixos, as conclusões ficam restritas aos tratamentos em teste, mas se os efeitos dos tratamentos são aleatórios, as conclusões devem ser estendidas para toda a população de onde os tratamentos foram amostrados. O pesquisador de posse dos resultados faz uso de um ramo da estatística chamado inferência estatística ou estatística indutiva, como se refere alguns autores, que faz uso das informações contidas numa amostra (ou experimento) para concluir para toda população.

O objetivo principal deste trabalho é apresentar os aspectos téoricos envolvidos na análise estatística de um experimento em blocos casualizados com efeitos aleatórios, os quais não são encontrados facilmente na literatura especializada e quando são encontrados carecem de clareza e justificativas teóricas. Além disso, é usado um exemplo real para ilustrar toda teoria, com interpretações adequadas ao problema proposto.

1.1 Princípios básicos da experimentação

Esta seção tem como objetivo apresentar alguns conceitos básicos da experimentação, que fundamenta a análise estatística dos dados provenientes de um planejamento de experimento com tratamento em blocos casualizados. Dentre outros, os tipos de delineamento experimental são: *inteiramente casualizado, em blocos ao acaso, em quadrados latinos*, etc. No delineamento experimental *inteiramente casualizado* estuda-se apenas duas fontes de variação, que são: tratamentos e erro, porém no delineamento experimental *em blocos completos ao acaso* são três fontes: tratamentos, blocos e erro. Os princípios básicos da experimentação são três: *Repetição, casualização* e *controle local*.

1.1.1 Repetição

Refere-se ao uso de mais de uma unidade experimental por tratamento. Assim, se em um experimento os tratamentos estiverem repetidos J vezes, diz-se que o ensa io ou experimento possui J repetições. Tal princípio tem por finalidade propiciar a obtenção de uma estimativa do resíduo experimental, melhorar a precisão do experimento e fazer com que o teste de hipóteses seja possível. Santos(2008, p.42).

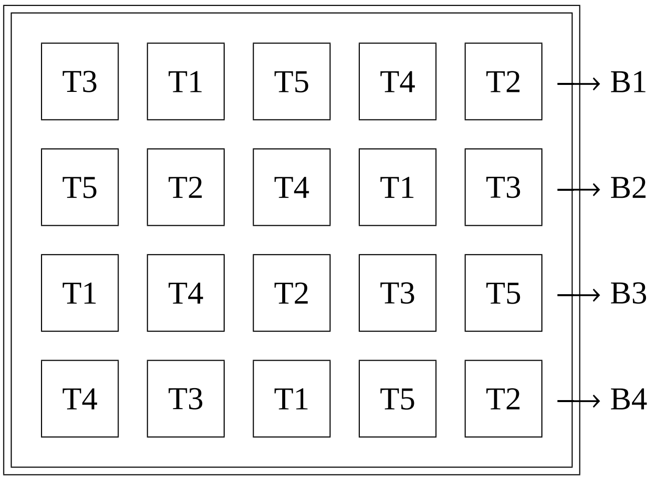
1.1.2 Casualização

A casualização consiste em se distribuir aleatoriamente os tratamentos nas parcelas, de modo que cada uma tenha a mesma chance de ocupar qualquer parcela na área experimental. A casualização assegura a validade da estimativa do erro experimental, pois permite uma distribuição do erro experimental. Vanderlei(2000, p.53).

1.1.3 Controle local

Controle local é usado quando as parcelas, antes de receberem os tratamentos, apresentam diferenças entre si. Dessa maneira, deve-se fazer o agrupamento das parcelas homogêneas em blocos, que têm por finalidade diminuir o erro experimental. Os critérios para agrupamentos das parcelas homogêneas em blocos podem ser: idade, sexo, peso textura, localização geográfica, topografia, etc.

Neste trabalho nos detemos ao estudo dos experimentos em blocos completos casualizados de efeitos aleatórios, cujo desenho desse tipo de ensaio no campo pode assumir várias configuraões. Dentre elas, considerando-se $I = 5$ tratamentos e $J = 4$ blocos, um possível desenho pode ser visto na figura a seguir.



Observe que:

1. cada bloco contém um conjunto completo dos tratamentos;

2. os tratamentos são distribuídos aleatoriamente dentro de cada bloco.

Os blocos controlam os efeitos de algumas variáveis externas, ou variáveis de perturbação. Este delineamento experimental talvez seja o tipo mais importante dentro da estatística experimental. Moore(2003,p.304).

Considerando que uma variável resposta Y esteja sendo mensurada, defina Y_{ij} como sendo o valor obtido da variável resposta e $y_{i.}$ é o valor obtido para o tratamento I e $\overline{y}_{i.}$ sua respectiva média. Para $y_{.j}$ o valor obtido para o bloco J e $\overline{y}_{.j}$ a média do bloco. E para $\overline{y}_{..}$ representa a média geral.

Tabela 1: Entrada de dados de um experimento em blocos completos casualizados.

Tratamentos	*Blocos*						*Total*	*Média*
(T_i)	1	2	$\cdots$	j	$\cdots$	J	$(y_{i.})$	$(m_i = \bar{y}_i)$
T_1	y_{11}	y_{12}	$\cdots$	y_{1j}	$\cdots$	y_{1J}	$y_{1.}$	$m_1 = \bar{y}_{1.}$
T_2	y_{21}	y_{22}	$\cdots$	y_{2j}	$\cdots$	y_{2J}	$y_{2.}$	$m_2 = \bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$
T_i	y_{i1}	y_{i2}	$\cdots$	y_{ij}	$\cdots$	y_{iJ}	$y_{i.}$	$m_i = \bar{y}_{i.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$
T_I	y_{I1}	y_{I2}	$\cdots$	y_{Ij}	$\cdots$	y_{IJ}	$y_{I.}$	$m_I = \bar{y}_{I.}$
Total	$y_{.1}$	$y_{.2}$	$\cdots$	$y_{.j}$	$\cdots$	$y_{.J}$	$y_{..}$	
Média	$\bar{y}_{.1}$	$\bar{y}_{.2}$	$\cdots$	$\bar{y}_{.j}$	$\cdots$	$\bar{y}_{.J}$		$m = \bar{y}_{..}$

1.2 Conceitos do plano experimental

1.2.1 Experimento

É um processo racional e sistematizado, é planejado para descobrir ou obter novos fatos, aceitar ou rejeitar hipóteses pesquisadas ou obtidas anteriormente ou em vigor. Experimento, também denominado de ensaio, é uma operação na qual se procura avaliar o efeito de vários tratamentos em conjunto, comparando-os em seguida baseado numa hipótese estatística. Considerando um teste estatístico ideal ou adequado é definido também como uma experiência que se aplica a causas controladas. Os blocos, são considerados em um experimeto fator que objetiva tornar um ambiente heterogêneo em homogêneo, controlando o ambiente ou material experimental afim de que não haja influência de outras variáveis ou variáveis de perturbação, como a luz, ventilação a temperatura, etc.

1.2.2 Tratamento

Tratamento é uma condição imposta ou objeto que se deseja medir ou avaliar em um experimento. A nível de pesquisa podemos classificar essa condição como delimitação do problema ou objetivo geral. Nos experimentos, em geral, é utilizado uma técnica para o planejamento do experimento, que por sua vez vai depender da natureza do problema a ser estudado. Os experimentos podem apresentar mais de um efeito de tratamento, os efeitos de tratamentos se dividem em dois tipos que são, efeitos de tratamentos aleatórios e efeitos de tratamentos fixos. Nos efeitos de tratamentos fixo o pesquisador, antes de iniciar a aplicação do experimento especifica ou escolhe o tratamento que irá comparar, já nos tratamentos com efeitos aleatórios, os tratamentos a ser comparados são sorteados. O pesquisador aplica as técnicas de amostragem aleatórias, para escolher os tratamentos que irá analisar. Os tratamentos se dividem em dois tipos, tratamento qualitativo e tratamento quantitativo. Os tratamentos que podem ser postos numa ordem, são chamados de tratamentos quantitativo. Por outro lado os tratamentos que não permitem uma ordem são chamados de tratamentos qualitativos. cada tipo de tratamento também pode ser chamado de fator. Quando em um experimento estamos interessados em estudar apenas um tipo de variável aleatória independente dizemos que possuimos apenas um fator. O tipo de tratamento tem importância significativa na forma como os dados serão analisados, influenciando nas devidas conclusões posteriores. Os tratamentos também são chamados de variáveis aleatórias independentes.

1.2.3 Testemunha

É um tratamento padrão, nele não se altera as condições naturais ou arbitrária, é a base para a comparação do estudo em que estamos realizando, seja doses de nutrientes, tipos de medicamentos, composição de uma peça de automóvel, etc. A importância da variável testemunha é tanto quanto as demais incluidas no experimento, deve ter as mesmas considerações, contínuas ou discretas,ordinal ou nominal qualitativas ou quantitativas. A variável testemunha é um ponto de partida e segue uma sequencia dentro da fonte de variação.

1.2.4 Resíduo experimental

Segundo Santos(2008, p.39) é a medida das variações existentes entre os dados ou observações que se apresentam nas unidades experimentais que recebem os tratamentos iguais, ou seja, é a causa da variação que reflete os efeitos do acaso. Classificação dos erros experimentais.

- Inerentes à variabilidade nas unidades experimentais, nas quais os tratamentos são aplicados, portanto é caracteristica destas unidades produzirem resultados diferentes quando submetidas ao mesmo tratamento. Logo, em outras palavras, dificilmente os valores observados a cada repetição dos tratamentos serão iguais.

- Ausência de uniformidade na condução física do experimento, isto é, quando há falhas na padronização da técnica aplicada.

1.2.5 Parcela

Segundo Vanderlei(2000, p.2), parcela é a unidade em que é feita a aplicação casualizada do tratamento , de modo a fornecer os dados experimentais que deverão refletir seu efeito. Em outras palavras, é a menor porção do material experimental onde os tratamentos são avaliados. Por exemplo, uma parcela pode ser: uma única planta ou um grupo delas, uma área de terreno com plantas, um lote de sementes, um vaso de barro (argila), uma parte da copa de uma árvore ou copa inteira, uma caixa de madeira, uma placa de petri, etc.

1.2.6 Considerações Gerais do Delineamento em Blocos Casualizados

O delineamento em blocos casualizados, também denominado de delineamento em blocos completos casualizados, se constitui no delineamento estatistico mais utilizado na pesquisa agronômica devido a sua simplicidade, flexibilidade e alta precisão. Os experimentos instalados de acordo com este delineamento são denominados de experimentos em blocos casualizados. Os experimentos em blocos casualizados levam em consideração os três princípios básicos da experimentação: repetição, casualização e controle local, contudo o controle local é usado na sua forma mais simples possível e é representado pelos blocos. Vanderlei(200, p.141).

1.2.6.1 Vantagens na instalação de um Delineamento em Blocos Casualizados

- A formação de blocos permite eliminar, das comparações entre tratamentos e da variação residual, a parte da variação devida à heterogeneidade do material experimental atribuível às diferenças entre os blocos.

- Em tese, o delineamento em blocos ao acaso não limita o número de repetições dos tratamentos.

- A análise estatística de experiemntos em blocos casualizados é simples. A perda de um ou mais tratamento e até de um bloco não ocasiona complicações para a análise estatística.

- As parcelas perdidas são facilmente estimadas. Santos(2008, p.94).

- Apresenta um número razoável de graus de liberdade para o resíduo - Ele se encontra numa faixa intermediária entre o delineamento inteiramente casualizado, que apresenta um maior número de graus de liberdade para o resíduo e o delineamento em quadro latino. Sabe-se que quanto maior o número de graus de liberdade para o resíduo, sensibilidade terá os testes de hipóteses para detectar diferença significativa entre os tratamentos avaliados, além de proporcionar maior precisão experimental. Portanto, o delineamento em blocos casualizados apresenta essa vantagem em relação ao delineamento em quadrado latino. Vanderlei(2000, p.143).

1.2.6.2 Desvantagens na instalação de um Delineamento em Blocos Casualizados

- Quando o número de tratamentos por bloco aumenta grandemente, a eficiência deste delineamento se reduz muito, devido à perda de homogeneidade dos mesmos dentro dos blocos.
- Quando o número de parcelas perdidas é elevado, este delineamento é menos conviniente que o inteiramente casualizado.
- Exige que os tratamentos tenham o mesmo número de repetições.

1.3 Componentes de variância no modelo aleatório

1.3.1 Definição

Existem vários métodos de estimação de componentes de variância. Os mais antigos são baseados no método dos momentos, entre os quais se incluem o método da análise de variância ANOVA (Fisher, 1918). Segundo Barbin(1993, p.07) Componentes de variância são variâncias associadas aos efeitos aleatórios de um modelo matemático. Entendemos como efeito aleatório como sendo aqueles representativos de tratamentos de uma amostra oriunda de uma população. Por exemplo: Se de um rebanho tiramos uma amostra de animais, estes animais entrarão no experimento representando todo o rebanho. Geralmente, blocos também são considerados de efeito aleatório pois que se pretende que os resultados experimentais sejam válidos para todo o local. De modo análogo os locais nos grupos de experimentos são considerados de efeito aleatório, já que se pretende estender os resultados para toda uma região. Existem outros dois tipos de modelos que são envolvidos na análise dos experimentos, os modelos de efeito fixo e efeito misto, estes não serão estudados neste trabalho.

2 Desenvolvimento teórico

2.1 Efeitos fixos, efeitos aleatórios e hipóteses de interesse

O principal objetivo deste trabalho é apresentar a fundamentação Estatística-Matemética que dar suporte a análise de um experimento em blocos casualizados com efeitos aleatórios. Assim sendo, achou-se conveniente apresentar, neste espaço, uma discussão sobre o que caracteriza um *fator de efeitos fixos* e um *fator de efeitos aleatórios*.

Existem duas maneiras pelas quais(os tratamentos) são escolhidos para o experimento:

1. Quando (os tratamentos) são escolhidos (fixados) pelo pesquisador. Nesse caso, o interesse será estimar e contrastar hipóteses sobre os efeitos dos tratamentos, τ_i. Nestas circunstâncias, as conclusões serão válidas apenas para os tratamentos considerados, não podendo ser estendidas a outros tratamentos que não foram considerados no experimento. Sendo assim, diz-se que o modelo é de *efeitos fixos*;

2. Quuando (os tratamentos) são sorteados aleatoriamente de uma população de níveis (de tratamentos). Nessa situação as conclusões, baseadas numa amostra de tratamentos, serão válidas para toda população de tratamentos, tendo eles sido considerados explicitamente no experimento ou não. Dessa forma, os τ_i são variáveis aleatórias e o interesse será estimar e contrastar hipóteses sobre a variabilidade dos τ_i. Assim sendo, diz-se que o modelo é de *efeitos aleatórios* ou de *componentes da variância*.

Neste trabalho, considera-se o experimento em blocos casualizados de efeitos aleatórios, isto é, admite-se que os efeitos dos blocos, os efeitos dos tratamentos e os efeitos do acaso são aleatórios. Nestas circunstâncias, as conclusões acerca da variabilidade dos efeitos dos tratamentos e da variabilidade dos efeitos dos blocos serão estendidas para toda população da qual foi retirada a amostra de tratamentos e para toda população de onde foi retirada a amostra dos blocos.

Formalmente, o procedimento estatístico para análise dos dados de um experimento com estas características é baseado na contrastação das hipóteses:

$$H_0^{(\tau)} : \sigma_\tau^2 = 0 \quad \text{vs} \quad H_1^{(\tau)} : \sigma_\tau^2 > 0,$$

$$H_0^{(\beta)} : \sigma_\beta^2 = 0 \quad \text{vs} \quad H_1^{(\beta)} : \sigma_\beta^2 > 0,$$

seguido da estimação dos componentes da variância e da construção de interalos de confiaça para eles. Nas seções seguintes, serão apresentadas as justificativas teóricas para o cumprimento destas finalidades.

2.2 O modelo estatístico

O modelo Estatístico linear apropriado para descrever as observações de um experimento em blocos ao acaso é denotado por:

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \qquad \begin{cases} i = 1, 2, \cdots, I \\ j = 1, 2, \cdots, J, \end{cases} \tag{2.1}$$

em que,

- y_{ij} é a observação obtida da unidade experimental que recebeu o i-ésimo tratamento no j-ésimo bloco;

- μ é a média geral;

- τ_i é o efeito do i-ésimo tratamento sobre a variável resposta, considerado aleatório e distribuido segundo uma normal com média *zero* e variância σ_τ^2. Supõe-se, também, que os τ_i's são independentemente distribuídos, isto é,

$$\tau_i \overset{iid}{\sim} N\left(0;\ \sigma_\tau^2\right), \quad i = 1, \cdots, I.$$

- β_j é o efeito do j-ésimo bloco sobre a variável resposta, também considerado aleatório e distribuído segundo uma normal com média *zero* e variância σ_β^2. Ademais, os β_j's são distribuídos independentemente um do outro, ou seja,

$$\beta_j \overset{iid}{\sim} N(0;\ \sigma_\beta^2), \quad j = 1, \cdots, J.$$

- ϵ_{ij} é o erro experimental o qual pode ser entendido como o efeito de outras variáveis de pertubação não consideradas no modelo. Além disso, considera-se que os erros ϵ_{ij}

são variáveis aleatórias independentes e identicamente distribuidas como uma normal de média *zero* e variância σ^2, a qual denota-se por: $\epsilon_{ij} \stackrel{iid}{\sim} N\Big(0;\sigma^2\Big)$.

Adicionalmente, admite-se que os τ_i's são independentemente distribuídos dos β_j's e que, tantos os τ_i's quanto os β_j's são independentemente distribuídos dos ϵ_{ij}'s.

Considerando-se as caractrísticas atribuídas aos termos no modelo (2.1), verifica-se que:

$$E(\mu) = \mu; \quad E(\mu^2) = \mu^2; \quad E(\tau_i) = 0; \quad E(\tau_i^2) = \sigma_\tau^2;$$

$$E(\beta_j) = 0; \quad E(\beta_j^2) = \sigma_\beta^2; \quad E(\epsilon_{ij}) = 0; \quad E(\epsilon_{ij}^2) = \sigma^2;$$

$$E(\tau_i\tau_{i'}) = E(\tau_i)E(\tau_{i'}) = 0, \ \forall\ i \neq i'; \quad E(\beta_j\beta_{j'}) = E(\beta_j)E(\beta_{j'}) = 0, \ \forall\ j \neq j';$$

$$E(\tau_i\beta_j) = E(\tau_i)E(\beta_j) = 0; \quad E(\tau_i\epsilon_{ij}) = E(\tau_i)E(\epsilon_{ij}) = 0; \quad E(\beta_j\epsilon_{ij}) = E(\beta_j)E(\epsilon_{ij}) = 0.$$

2.3 Estimação dos efeitos do modelo

Para encontrar os estimadores dos efeitos do modelo (2.1), utilizou-se o método da máxima verossimilhança, o qual consiste de encontrar os valores dos componene do vetor de efeitos, θ, onde $\theta' = (\mu\ \tau_i\ \beta_j\ \sigma^2)$, que maximixam a função de verossimilhança, $L(\theta)$, ou seja, sendo

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \quad \Longrightarrow \quad \epsilon_{ij} = y_{ij} - \mu - \tau_i - \beta_j$$

e sabendo-se que

$$\epsilon_{ij} \stackrel{iid}{\sim} N(0;\ \sigma^2) \quad \Longrightarrow \quad f_\epsilon(\epsilon_{ij}|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_{ij}-\mu-\tau_i-\beta_j)^2},$$

então,

$$\begin{aligned} L(\theta) &= L(\mu,\tau_i,\beta_j,\sigma^2|\epsilon_{11},\cdots,\epsilon_{IJ}) = \prod_{i=1}^{I}\prod_{j=1}^{J} f_\epsilon(\epsilon_{ij}) \\ &= (2\pi\sigma^2)^{-1/2}e^{-\frac{1}{2\sigma^2}(y_{11}-\mu-\tau_1-\beta_1)^2}.\cdots.(2\pi\sigma^2)^{-1/2}e^{-\frac{1}{2\sigma^2}(y_{IJ}-\mu-\tau_I-\beta_J)^2} \\ &= (2\pi\sigma^2)^{-IJ/2}e^{-\frac{1}{2\sigma^2}\sum\limits_{i,j}(y_{ij}-\mu-\tau_i-\beta_j)^2}. \end{aligned}$$

O logarítmo da função de verossimilhança, fica.

$$\ell(\theta) = \ln L(\theta) = -\frac{IJ}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{ij}(y_{ij} - \mu - \tau_i - \beta_j)^2. \tag{2.2}$$

O método da máxima verossimilhança consiste em encontrar o estimador de θ, isto é, $\hat{\theta}$, de modo que maximaza a função de verossimilhança. Como $\ell(\theta)$ é uma função crescente de θ, tem-se que $L(\theta)$ será máxima para o mesmo $\hat{\theta}$ que maximiza $\ell(\theta)$. Para isto deriva-se parcialmente $\ell(\theta)$ em relação a cada componente do vetor θ, igualando-se a zero e explicita-se cada componente específico, ou seja:

$$\begin{aligned}
\frac{\partial \ell(\theta)}{\partial \mu} &= -\frac{2}{\sigma^2}\sum_{i,j}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j)(-1) = 0 \\
&\therefore \sum_{ij}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j) = 0, \quad \text{ou} \\
&IJ\hat{\mu} + J\sum_{i=1}^{I}\hat{\tau}_i + I\sum_{j=1}^{J}\hat{\beta}_j = y_{..}
\end{aligned} \tag{2.3}$$

de modo análogo, tem-se:

$$\begin{aligned}
\frac{\partial \ell(\theta)}{\partial \tau_i} &= -\frac{2}{2\sigma^2}\sum_{j}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j)(-1) = 0 \\
&\therefore \sum_{j}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j) = 0, \quad \text{ou} \\
&J\hat{\mu} + J\hat{\tau}_i + \sum_{j=1}^{J}\hat{\beta}_j = y_{i.},
\end{aligned} \tag{2.4}$$

$$\begin{aligned}
\frac{\partial \ell(\theta)}{\partial \beta_j} &= -\frac{2}{2\sigma^2}\sum_{i}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j)(-1) = 0 \\
&\therefore \sum_{i}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j) = 0, \quad \text{ou} \\
&I\hat{\mu} + \sum_{i=1}^{I}\hat{\tau}_i + I\hat{\beta}_j = y_{.j}
\end{aligned} \tag{2.5}$$

e, finalmente

$$\begin{aligned}
\frac{\partial \ell(\theta)}{\partial \sigma^2} &= -\frac{IJ}{2}.\frac{1}{2\pi\sigma^2}.2\pi + \frac{1}{2}.\frac{1}{\sigma^4}\sum_{i,j}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j)^2 = 0 \\
&\therefore -\frac{IJ}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\sum_{i,j}(y_{ij} - \hat{\mu} - \hat{\tau}_i - \hat{\beta}_j)^2 = 0.
\end{aligned} \tag{2.6}$$

Multiplicando-se os dois lados da equação (2.6) por $2\hat{\sigma}^2$ e fazendo-se algumas operações algébricas simples, obtém-se:

$$\hat{\sigma}^2 = \frac{1}{IJ}\sum_{i,j}(y_{ij} - \hat{y}_{ij})^2, \tag{2.7}$$

isto é, o estimador da variância do erro é a média dos quadrados dos resíduos.

As equações (2.3), (2.4) e (2.5) formam o sistema de equações normais (S.E.N). Isto é,

$$\begin{cases} IJ\hat{\mu} + J\sum_i \hat{\tau}_i + I\sum_j \hat{\beta}_j = y_{..} \quad (I) \\ \\ J\hat{\mu} + J\hat{\tau}_i + \sum_j \hat{\beta}_j = y_{i.} \quad (II) \\ \\ I\hat{\mu} + \sum_i \hat{\tau}_i + I\hat{\beta}_j = y_{.j} \quad (III) \end{cases} \tag{2.8}$$

O sistema de equações normais (S.E.N) em (2.8) é inconsistente, uma vez que é formado por *três* equações e $(I + J + 1)$ incógnitas (estimadores) e $(I,\ J > 1)$. Para resolvê-lo, em geral, admite-se as restrições $\sum_i \hat{\tau}_i = 0$ e $\sum_j \hat{\beta}_j = 0$, conhecidas na literatura como *restrições na solução*.

Assim sendo, os estimadores de μ, τ_i e β_j, são obtidos facilmente. Ou seja,

$$\hat{\mu} = \bar{y}_{..}, \quad \hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..} \quad \text{e} \quad \hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..}. \tag{2.9}$$

Em palavras: o estimador da média geral, μ, é a média dos valores observados no experimento, $\bar{y}_{..}$, o estimador do efeito do i-ésimo tratamento sobre a variável resposta é a diferença entre a média dos valores observados no tratamento i e a média geral e o estimador do efeito do j-ésimo bloco sobre a variável resposta, é a diferença entre a média das observações no bloco j, e a média geral.

Os estimadores de y_{ij} e ϵ_{ij}, são dados, respectivamente, por:

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j \qquad \text{e} \qquad \hat{\epsilon}_{ij} = e_{ij} = y_{ij} - \hat{y}_{ij}.$$

Substituindo-se os resultados encontrados em (2.9), obtém-se,

$$\hat{y}_{ij} = \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..} \qquad \text{e} \qquad e_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..} \tag{2.10}$$

Vale salientar que os reíduos e_{ij} são elementos importantes a serem utilizados para comprovar as suposições impostas aos termos no modelo: *Aditividade*, *Normalidade*, *independência e homocedasticidade*.

2.4 Decomposição da variabilidade total

Cada observação experimental pode ser escrita pela identidade:

$$y_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_j + \hat{\epsilon}_{ij}$$

ou

$$y_{ij} = \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) \tag{2.11}$$

ou, ainda

$$(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) = (y_{ij} - \bar{y}_{..}) - (\bar{y}_{i.} - \bar{y}_{..}) - (\bar{y}_{.j} - \bar{y}_{..}). \tag{2.12}$$

Elevando-se ao quadrado os dois lado da equação (2.12) e somando-se para todas as observações, após algumas operações algébricas simpes, obtém-se o seguinte resultado:

$$\underbrace{\sum_{i,j}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2}_{SQResiduos} = \underbrace{\sum_{i,j}(y_{ij} - \bar{y}_{..})^2}_{SQTotal} - \underbrace{J\sum_{i}(\bar{y}_{i.} - \bar{y}_{..})^2}_{SQTratamentos} - \underbrace{I\sum_{j}(\bar{y}_{.j} - \bar{y}_{..})^2}_{SQBlocos}. \tag{2.13}$$

É fácil demonstrar que os termos da equação (2.13) podem ser escritos da seguinte maneira:

$$SQResiduos = \underbrace{\left(\sum_{i,j} y_{ij}^2 - C\right)}_{SQTotal} - \underbrace{\left(\frac{1}{J}\sum_{i=1}^{I} y_{i.}^2 - C\right)}_{SQTratamentos} - \underbrace{\left(\frac{1}{I}\sum_{j=1}^{J} y_{.j}^2 - C\right)}_{SQBlocos}, \quad \text{onde} \quad C = \frac{(y_{..})^2}{IJ}. \tag{2.14}$$

Dessa forma, observa-se que a variabilidade total é decomposta em três partes, a saber:

$$SQTotal = SQTratamentos + SQBlocos + SQResiduos.$$

Na prática, a variabilidade do acaso ou variação residual é calculada por diferença, conforme a equação (2.14), isto é,

$$SQResiduo = SQTotal - SQTratamentos - SQBlocos.$$

2.5 Distribuições de probabilidade dos estimadores dos efeitos dos componentes do modelo e das somas dos quadrados

1. **Distribuições de probabilidade de $\hat{\tau}_i$ e da $SQTrat$.**

Já é sabido que:

$$\begin{aligned}
\hat{\tau}_i &= \overline{y}_{i.} - \overline{y}_{..} \text{ é uma combinação linear de normais e, portanto, } \hat{\tau}_i \text{ é normal.}\\
&= \frac{1}{J}\sum_j y_{ij} - \frac{1}{IJ}\sum_{i,j} y_{ij}\\
&= \frac{1}{J}\left(J\mu + J\tau_i + \sum_j \beta_i + \sum_j \epsilon_{ij}\right) - \frac{1}{IJ}\left(IJ\mu + J\sum_i \tau_i + I\sum_j \beta_j + \sum_{i,j}\epsilon_{ij}\right)\\
&= \tau_i - \frac{1}{I}\sum_i \tau_i + \frac{1}{J}\sum_i \epsilon_{ij} - \frac{1}{IJ}\sum_{i,j}\epsilon_{ij}. \qquad (2.15)
\end{aligned}$$

Dessa forma, tem-se que

$$E(\hat{\tau}_i) = 0$$

e

$$\begin{aligned}
Var(\hat{\tau}_i) &= E[\hat{\tau}_i - E(\hat{\tau}_i)]^2\\
&= E\left[\tau_i - \frac{1}{I}\sum_i \tau_i + \frac{1}{J}\sum_i \epsilon_{ij} - \frac{1}{IJ}\sum_{i,j}\epsilon_{ij}\right]^2\\
&= E\left[\tau_i^2 + \frac{1}{I^2}\left(\sum_i \tau_i\right)^2 + \frac{1}{J^2}\left(\sum_j \epsilon_{ij}\right)^2 + \frac{1}{I^2J^2}\left(\sum_{i,j}\epsilon_{ij}\right)^2 - \frac{2}{I}\tau_i\sum_i \tau_i\right.\\
&\quad \left. -\frac{2}{IJ^2}\left(\sum_j \epsilon_{ij}\right)\left(\sum_{i,j}\epsilon_{ij}\right) + odp = \text{outros duplos produtos}\right]\\
&= \frac{(I-1)}{I}\frac{(\sigma^2 + J\sigma^2_\tau)}{J}.
\end{aligned}$$

Assim sendo, tem-se que

$$\hat{\tau}_i = \overline{y}_{i.} - \overline{y}_{..} \sim N\left(0;\ \frac{(I-1)}{I}\frac{(\sigma^2 + J\sigma^2_\tau)}{J}\right) \qquad (2.16)$$

e segue que

$$\frac{\overline{y}_{i.} - \overline{y}_{..}}{\sqrt{\frac{(I-1)}{I}\frac{(\sigma^2+J\sigma^2_\tau)}{J}}} \sim N(0;\ 1) \Longrightarrow \frac{(\overline{y}_{i.} - \overline{y}_{..})^2}{\frac{(I-1)}{I}\frac{(\sigma^2+J\sigma^2_\tau)}{J}} \sim \chi^2_{(1)}$$

ou

$$\frac{J(\overline{y}_{i.} - \overline{y}_{..})^2}{\sigma^2 + J\sigma^2_\tau} \sim \frac{(I-1)}{I}\chi^2_{(1)}.$$

Como $(\overline{y}_{1.} - \overline{y}_{..}), (\overline{y}_{2.} - \overline{y}_{..}), \cdots, (\overline{y}_{I.} - \overline{y}_{..})$ são variáveis aleatórias independentes e estão sujeita a restrição $\sum_i \hat{\tau}_i = \sum_i (\overline{y}_{i.} - \overline{y}_{..}) = 0$, então

$$\frac{J\sum_{i=1}^{I}(\overline{y}_{i.} - \overline{y}_{..})^2}{\sigma^2 + J\sigma^2_\tau} = \frac{SQTrat.}{\sigma^2 + J\sigma^2_\tau} \sim (I-1)\chi^2_{(1)} = \chi^2_{(I-1)}. \qquad (2.17)$$

Portanto, sob a hipótese, $H_0 : \sigma^2_\tau = 0$, a estatística $\frac{SQTrat.}{\sigma^2}$ segue uma distribuição de qui-quadrado central com $(I-1)$ graus de liberdade, a qual será denotada por:

$$\frac{SQTrat.}{\sigma^2} \sim \chi^2_{(I-1)}. \tag{2.18}$$

2. Distribuições de probabilidade de $\hat{\beta}_j$ e da $SQBlocos$

De acordo com (2.9), tem-se que:

$$\begin{aligned}
\hat{\beta}_j &= \bar{y}_{.j} - \bar{y}_{..} \text{ é uma combinação linear de normais e, portanto, } \hat{\beta}_j \text{ é normal.}\\
&= \frac{1}{I}\sum_i y_{ij} - \frac{1}{IJ}\sum_{i,j} y_{ij}\\
&= \frac{1}{I}\left(I\mu + \sum_i \tau_i + I\beta_j + \sum_i \epsilon_{ij}\right) - \frac{1}{IJ}\left(IJ\mu + J\sum_i \tau_i + I\sum_j \beta_j + \sum_{i,j}\epsilon_{ij}\right)\\
&= \beta_j - \frac{1}{J}\sum_j \beta_j + \frac{1}{I}\sum_i \epsilon_{ij} - \frac{1}{IJ}\sum_{i,j}\epsilon_{ij}. \qquad (2.19)
\end{aligned}$$

Daí, obtém-se:

$$E(\hat{\beta}_j) = 0$$

e

$$\begin{aligned}
Var(\hat{\beta}_j) &= E[\hat{\beta}_j - E(\hat{\beta}_j)]^2\\
&= E\left[\beta_j - \frac{1}{J}\sum_j \beta_j + \frac{1}{I}\sum_i \epsilon_{ij} - \frac{1}{IJ}\sum_{i,j}\epsilon_{ij}\right]^2\\
&= E\left[\beta_j^2 + \frac{1}{J^2}\left(\sum_j \beta_j\right)^2 + \frac{1}{I^2}\left(\sum_i \epsilon_{ij}\right)^2 + \frac{1}{I^2J^2}\left(\sum_{i,j}\epsilon_{ij}\right)^2 - \frac{2}{J}\beta_j\sum_j \beta_j\right.\\
&\qquad \left. -\frac{2}{I^2J}\left(\sum_i \epsilon_{ij}\right)\left(\sum_{i,j}\epsilon_{ij}\right) + odp\right]\\
&= \sigma^2_\beta + \frac{1}{J}\sigma^2_\beta + \frac{1}{I}\sigma^2 + \frac{1}{IJ}\sigma^2 - \frac{2}{J}\sigma^2_\beta - \frac{2}{IJ}\sigma^2 + 0 + \cdots + 0\\
&= \frac{(J-1)}{J}\frac{(\sigma^2 + I\sigma^2_\beta)}{I}. \qquad (2.20)
\end{aligned}$$

Portanto,

$$\hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..} \sim N\left(0;\ \frac{(J-1)}{J}\frac{(\sigma^2 + I\sigma^2_\beta)}{I}\right) \Longrightarrow \frac{\bar{y}_{.j} - \bar{y}_{..}}{\sqrt{\frac{(J-1)}{J}\frac{(\sigma^2+I\sigma^2_\beta)}{I}}} \sim N(0;\ 1) \tag{2.21}$$

e

$$\frac{(\bar{y}_{.j}-\bar{y}_{..})^2}{\frac{(J-1)}{J}\frac{(\sigma^2+I\sigma^2_\beta)}{I}} \sim \chi^2_{(1)} \Longrightarrow \frac{I(\bar{y}_{.j}-\bar{y}_{..})^2}{\sigma^2+I\sigma^2_\beta} \sim \frac{(J-1)}{J}\chi^2_{(1)}.$$

Como $(\bar{y}_{.1}-\bar{y}_{..}), (\bar{y}_{.2}-\bar{y}_{..}), \cdots, (\bar{y}_{.J}-\bar{y}_{..})$, são variáveis aleatórias independentes e sujeitas à restrição, $\sum_j \hat{\beta}_j = \sum_j (\bar{y}_{.j}-\bar{y}_{..}) = 0$, então,

$$\frac{I\sum_j(\bar{y}_{.j}-\bar{y}_{..})^2}{\sigma^2+I\sigma^2_\beta} = \frac{SQBlocos}{\sigma^2+I\sigma^2_\beta} \sim (J-1)\chi^2_{(1)} = \chi^2_{(J-1)} \tag{2.22}$$

Além disso, sob $H_0 : \sigma^2_\beta = 0$, a estatística $\frac{SQBlocos}{\sigma^2}$ segue uma distribuição de qui-quadrado central com $(J-1)$ grus de liberdade e será denotada por:

$$\frac{SQBlocos}{\sigma^2} \sim \chi^2_{(J-1)}. \tag{2.23}$$

3. Distribuições de probabilidade de $\hat{\epsilon}_{ij}$ e da $SQResiduos$

De acordo com (2.10), os resíduos são dados por:

$$\begin{aligned}
\hat{\epsilon}_{ij} &= y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..} \text{ é uma combinação linear de normais e, portanto, } \hat{\epsilon}_{ij} \text{ é normal.}\\
&= \mu+\tau_i+\beta_j+\epsilon_{ij}-\frac{1}{J}\left(J\mu+J\tau_i+\sum_j\beta_j+\sum_j\epsilon_{ij}\right)\\
&\quad -\frac{1}{I}\left(I\mu+\sum_i\tau_i+I\beta_j+\sum_i\epsilon_{ij}\right)+\frac{1}{IJ}\left(IJ\mu+J\sum_i\tau_i+I\sum_j\beta_j+\sum_{i,j}\epsilon_{ij}\right)\\
&= \epsilon_{ij}-\frac{1}{I}\sum_i\epsilon_{ij}-\frac{1}{J}\sum_j\epsilon_{ij}+\frac{1}{IJ}\sum_{i,j}\epsilon_{ij}. \qquad (2.24)
\end{aligned}$$

Sendo assim, tem-se que:

$$E(\hat{\epsilon}_{ij}) = 0$$

e

$$
\begin{aligned}
Var(\hat{\epsilon}_{ij}) &= E[\hat{\epsilon}_{ij} - E(\hat{\epsilon}_{ij})]^2 = E\left[\epsilon_{ij} - \frac{1}{I}\sum_i \epsilon_{ij} - \frac{1}{J}\sum_j \epsilon_{ij} + \frac{1}{IJ}\sum_{i,j}\epsilon_{ij}\right]^2 \\
&= E\left[\epsilon_{ij}^2 + \frac{1}{I^2}\left(\sum_i \epsilon_{ij}\right)^2 + \frac{1}{J^2}\left(\sum_j \epsilon_{ij}\right)^2 + \frac{1}{I^2J^2}\left(\sum_{i,j}\epsilon_{ij}\right)^2 - \frac{2}{I}\epsilon_{ij}\sum_i \epsilon_{ij}\right. \\
&\quad -\frac{2}{J}\epsilon_{ij}\sum_j \epsilon_{ij} + \frac{2}{IJ}\epsilon_{ij}\sum_{i,j}\epsilon_{ij} + \frac{2}{IJ}\left(\sum_i \epsilon_{ij}\right)\left(\sum_j \epsilon_{ij}\right) - \frac{2}{I^2J}\left(\sum_i \epsilon_{ij}\right)\left(\sum_{ij}\epsilon_{ij}\right) \\
&\quad \left. -\frac{2}{IJ^2}\left(\sum_j \epsilon_{ij}\right)\left(\sum_{i,j}\epsilon_{ij}\right)\right] \\
&= \sigma^2 + \frac{1}{I}\sigma^2 + \frac{1}{J}\sigma^2 + \frac{1}{IJ}\sigma^2 - \frac{2}{I}\sigma^2 - \frac{2}{J}\sigma^2 + \frac{2}{IJ}\sigma^2 + \frac{2}{IJ}\sigma^2 - \frac{2}{IJ}\sigma^2 - \frac{2}{IJ}\sigma^2 \\
&= \frac{(I-1)(J-1)}{IJ}\sigma^2. \qquad (2.25)
\end{aligned}
$$

Portanto,

$$
\hat{\epsilon}_{ij} = y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..} \;\sim\; N\left(0;\ \frac{(I-1)(J-1)}{IJ}\sigma^2\right) \Longrightarrow \frac{y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..}}{\sqrt{\frac{(I-1)(J-1)}{IJ}\sigma^2}} \;\sim\; N(0;\ 1) \qquad (2.26)
$$

e

$$
\frac{(y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..})^2}{\frac{(I-1)(J-1)}{IJ}\sigma^2} \;\sim\; \chi^2_{(1)} \Longrightarrow \frac{(y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..})^2}{\sigma^2} \;\sim\; \frac{(I-1)(J-1)}{IJ}\chi^2_{(1)}.
$$

Assim sendo, dada a indepedência entre os resíduos, segue que

$$
\frac{\sum_{i,j}(y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..})^2}{\sigma^2} = \frac{SQResiduos}{\sigma^2} \;\sim\; (I-1)(J-1)\chi^2_{(1)} = \chi^2_{[(I-1)(J-1)]}. \qquad (2.27)
$$

Isto é, a estatística, $\frac{SQResiduos}{\sigma^2}$, segue uma distribuição de qui-quadrado central com $(I-1)(J-1)$ graus de liberdade.

4. Distribuição de probabilidade da $SQTotal$

Também, é necessário conhecer a distribuição de probabilidade da $SQTotal$ sob a hipótese de que não existe variabilidade entre os efitos dos tratamentos e nem dos efeitos dos blocos. Isto é, sob a hipótese de que as observações podem ser escritas pelo modelo $y_{ij} = \mu + \epsilon_{ij}$, com $i = 1, \cdots, I$ e $j = 1, \cdots, J$.

Tomando-se as diferenças $y_{ij} - \bar{y}_{..}$, vem:

$$\begin{aligned} y_{ij} - \bar{y}_{..} &= \mu + \epsilon_{ij} - \frac{1}{IJ}\left(IJ\mu + \sum_{i,j}\epsilon_{ij}\right) \\ &= \epsilon_{ij} - \frac{1}{IJ}\sum_{i,j}\epsilon_{ij} \end{aligned}$$

Daí, obtém-se os seguintes resultados:

$$E(y_{ij} - \bar{y}_{..}) = 0$$

e

$$\begin{aligned} Var(y_{ij} - \bar{y}_{..}) &= E[y_{ij} - \bar{y}_{..} - E(y_{ij} - \bar{y}_{..})]^2 = E\left[\epsilon_{ij} - \frac{1}{IJ}\sum_{i,j}\epsilon_{ij}\right]^2 \\ &= E\left[\epsilon_{ij}^2 + \frac{1}{I^2J^2}\left(\sum_{i,j}\epsilon_{ij}\right)^2 - \frac{2}{IJ}\epsilon_{ij}\sum_{i,j}\epsilon_{ij}\right] = \frac{(IJ-1)}{IJ}\sigma^2. \end{aligned}$$

Portanto, sob $H_0^{(\tau)} : \sigma_\tau^2 = 0$ e $H_0^{(\beta)} : \sigma_\beta^2 = 0$,

$$y_{ij} - \bar{y}_{..} \sim N\left(0;\ \frac{(IJ-1)}{IJ}\sigma^2\right) \Longrightarrow \frac{y_{ij} - \bar{y}_{..}}{\sqrt{\frac{(IJ-1)}{IJ}\sigma^2}} \sim N(0;\ 1)$$

e

$$\frac{(y_{ij} - \bar{y}_{..})^2}{\frac{(IJ-1)}{IJ}\sigma^2} \sim \chi^2_{(1)} \Longrightarrow \frac{(y_{ij} - \bar{y}_{..})^2}{\sigma^2} \sim \frac{(IJ-1)}{IJ}\chi^2_{(1)}.$$

E, admitindo a independência entre as diferenças, $(y_{11} - \bar{y}_{..}), (y_{12} - \bar{y}_{..}), \cdots, (y_{IJ} - \bar{y}_{..})$, então,

$$\frac{\sum_{i,j}(y_{ij} - \bar{y}_{..})^2}{\sigma^2} = \frac{SQTotal}{\sigma^2} \sim (IJ-1)\chi^2_{(1)} = \chi^2_{(IJ-1)}. \qquad (2.28)$$

Em palavras, sob $H_0^{(\tau)} : \sigma_\tau^2 = 0$ e $H_0^{(\beta)} : \sigma_\beta^2 = 0$, a estatística $\frac{SQTotal}{\sigma^2}$, segue uma distribuição de qui-quadrado central com $(IJ-1)$ graus de liberdade.

Na **Tabela** 2 estão apresentadas, de modo resumido, as distribuições de probabilidades dos estimadores dos efeitos do modelo e das somas dos quadrados.

Tabela 2: Distribuições de probabilidades dos estimadores dos efeitos do modelo e das somas dos quadrados

Distribuição do estimador	sob a(s) hipótese(s)	$\longrightarrow$	Distribuição da SQ sob $H_0^{(\cdot)}$
$\hat{\tau}_i \sim N\left(0;\ \frac{(I-1)}{I}\frac{(\sigma^2+J\sigma^2_\tau)}{J}\right)$	$H_0^{(\tau)}: \sigma^2_\tau = 0$	$\longrightarrow$	$\frac{SQTrat.}{\sigma^2} \overset{sob\ H_0^{(\tau)}}{\sim} \chi^2_{(I-1)}$
$\hat{\beta}_j \sim N\left(0;\ \frac{(J-1)}{J}\frac{(\sigma^2+I\sigma^2_\beta)}{I}\right)$	$H_0^{(\beta)}: \sigma^2_\beta = 0$	$\longrightarrow$	$\frac{SQBlocos}{\sigma^2} \overset{sob\ H_0^{(\beta)}}{\sim} \chi^2_{(J-1)}$
$\hat{\epsilon}_{ij} \sim N\left(0; \frac{(I-1)(J-1)}{IJ}\sigma^2\right)$	*Indep. de* $H_0^{(\tau)}$ e $H_0^{(\beta)}$	$\longrightarrow$	$\frac{SQRes.}{\sigma^2} \sim \chi^2_{[(I-1)(J-1)]}$
$y_{ij} - \bar{y}_{..} \sim N\left(0;\ \frac{(IJ-1)}{IJ}\sigma^2\right)$	$H_0^{(\tau)}$ e $H_0^{(\beta)}$	$\longrightarrow$	$\frac{SQTotal}{\sigma^2} \overset{sob\ H_0^{(\tau)}\ e\ H_0^{(\beta)}}{\sim} \chi^2_{(IJ-1)}$

2.6 Valores esperados das somas dos quadrados e dos quadrados médios

1. Valor esperado da soma de quadrados total

O valor esperado da $SQTotal$, considerando a expressão (2.13), é obtido do seguinte modo:

$$E(SQTotal) = E\left(\sum_{i,j} y_{ij}^2 - C\right) = E\left(\sum_{i,j} y_{ij}^2\right) - E(C).$$

Mas,

$$\begin{aligned} E\left(\sum_{i,j} y_{ij}^2\right) &= \sum_{i,j} E(\mu + \tau_i + \beta_j + \epsilon_{ij})^2 = \sum_{i,j} E(\mu^2 + \tau_i^2 + \beta_j^2 + \epsilon_{ij} + dp) \\ &= \sum_{i,j} (\mu^2 + \sigma^2_\tau + \sigma^2_\beta + \sigma^2 + 0) \\ &= IJ\mu^2 + IJ\sigma^2_\tau + IJ\sigma^2_\beta + IJ\sigma^2 \end{aligned} \tag{2.29}$$

e

$$\begin{aligned} E(C) &= \frac{1}{IJ} E\left(\sum_{i,j} y_{ij}\right)^2 = \frac{1}{IJ} E\left(IJ\mu + J\sum_i \tau_i + I\sum_j \beta_j + \sum_{i,j} \epsilon_{ij}\right)^2 \\ &= \frac{1}{IJ} E\left[J^2 I^2 \mu^2 + J^2\left(\sum_i \tau_i\right)^2 + I^2\left(\sum_j \beta_j\right)^2 + \left(\sum_{i,j} \epsilon_{ij}\right)^2 + dp\right] \\ &= IJ\mu^2 + J\sigma^2_\tau + I\sigma^2_\beta + \sigma^2 \end{aligned} \tag{2.30}$$

Subtraindo-se (2.30) de (2.29), obtem-se:

$$E(SQTotal) = J(I-1)\sigma_\tau^2 + I(J-1)\sigma_\beta^2 + (IJ-1)\sigma^2. \tag{2.31}$$

2. Valores esperados da soma de quadrados e do quadrado médio de tratamentos

$$E(SQTrat.) = E\left(\frac{1}{J}\sum_i y_{i.}^2 - C\right) = \frac{1}{J}E\left(\sum_i y_{i.}^2\right) - E(C).$$

Mas,

$$\begin{aligned}
\frac{1}{J}E\left(\sum_i y_{i.}^2\right) &= \frac{1}{J}\sum_i E\left(J\mu + J\tau_i + \sum_j \beta_j + \sum_j \epsilon_{ij}\right)^2\\
&= \frac{1}{J}\sum_i E\left[J^2\mu^2 + J^2\tau_i^2 + \left(\sum_j \beta_j\right)^2 + \left(\sum_j \epsilon_{ij}\right)^2 + dp\right]\\
&= IJ\mu^2 + IJ\sigma_\tau^2 + I\sigma_\beta^2 + I\sigma^2. \qquad (2.32)
\end{aligned}$$

Subtraindo-se (2.30) de (2.32), encontra-se:

$$E(SQTrat.) = J(I-1)\sigma_\tau^2 + (I-1)\sigma^2. \tag{2.33}$$

O valor esperado da $QMTrat.$, é obtido do seguinte modo:

$$E(QMTrat.) = E\left(\frac{SQTrat.}{I-1}\right) = \sigma^2 + J\sigma_\tau^2. \tag{2.34}$$

3. Valor esperado da soma de quadrados e do quadrado médio de blocos

$$E(SQBlocos) = E\left(\frac{1}{I}\sum_j y_{.j}^2 - C\right) = \frac{1}{I}E\left(\sum_j y_{.j}^2\right) - E(C).$$

Mas,

$$\begin{aligned}
\frac{1}{I}E\left(\sum_j y_{.j}^2\right) &= \frac{1}{I}E\left[\sum_j\left(I\mu + \sum_i \tau_i + I\beta_j + \sum_i \epsilon_{ij}\right)^2\right]\\
&= \frac{1}{I}\sum_j E\left[I^2\mu^2 + \left(\sum_i \tau_i\right)^2 + I^2\beta_j^2 + \left(\sum_i \epsilon_{ij}\right)^2 + dp\right]\\
&= IJ\mu^2 + J\sigma_\tau^2 + IJ\sigma_\beta^2 + J\sigma^2. \qquad (2.35)
\end{aligned}$$

Subtraindo-se (2.30) de (2.35), obtem-se:

$$E(SQBlocos) = I(J-1)\sigma_{\beta}^{2} + (J-1)\sigma^{2} \tag{2.36}$$

e o valor esperado do $QMBlocos$ é obtido da seguinte maneira:

$$E(QMBlocos) = E\left(\frac{SQBlocos}{J-1}\right) = \sigma^{2} + I\sigma_{\beta}^{2}. \tag{2.37}$$

4. Valor esperado da soma de quadrados e do quadrado médio dos resíduos

$$E(SQRes.) = E(SQTotal) - E(SQTrat.) - E(SQBlocos).$$

Tomando-se as expressões (2.33) e (2.36) e subtraindo-se da expressão (2.31), encontra-se o valor esperado da *SQResíduos*. Isto é,

$$E(SQRes.) = (I-1)(J-1)\sigma^{2} \tag{2.38}$$

O valor esperado do quadrado médio dos resíduos é dado por:

$$E(QMRes.) = E\left(\frac{SQRes.}{(I-1)(J-1)}\right) = \sigma^{2}. \tag{2.39}$$

Daí, conclui-se que o quadrado médio dos resíduos é um estimador não viciado para a variância do erro experimental, σ^2.

A **Tabela** 3 exibe um resumo dos resultados importantes obtidos nesta seção.

Tabela 3: Tabela da análise de variância com as esperaças dos quadrados médios

F. Variação	GL	SQ	QM	$E(QM)$	F
Tratamentos	$(I-1)$	$SQTrat.$	$QMTrat.$	$\sigma^2 + J\sigma^2_\tau$	$QMTrat./QMRes.$
Blocos	$(J-1)$	$SQBlocos$	$QMBlocos$	$\sigma^2 + J\sigma^2_\beta$	$QMBlocos/QMRes.$
Resíduos	$(I-1)(J-1)$	$SQRes.$	$QMRes.$	σ^2	-
Total	$IJ-1$	$SQTotal$	-	-	-

2.7 Análises estatísticas

1. De acordo com Rohatgi, Mood e Hogg, dentre outros, uma variável aleatória com distribuição F é dada pela razão entre duas variáveis aleatórias com distribuição de qui-quadrado independentes, divididas pelos seus respectivos graus de liberdade. Ou seja, se $U \sim \chi^2_{[\nu_1]}$, $V \sim \chi^2_{[\nu_2]}$ e U e V são variáveis aleatórias independente então, a estatística:

$$F = \frac{U/\nu_1}{V/\nu_2}$$

segue uma distribuição de probabilidade F com ν_1 graus de liberdade do numerador e ν_2 graus de liberdade do denominador, a qual é denotada por:

$$F = \frac{U/\nu_1}{V/\nu_2} \;\sim\; F_{[\nu_1;\,\nu_2]}.$$

2. Na seção 2.4, mostrou-se que a variabilidade total é decomposta em *três* partes independentes, a saber:

$$SQTotal = SQTrat. + SQBlocos + SQRes.$$

3. Na seção 2.5, demonstrou-se também, que:

$$\frac{SQTrat.}{\sigma^2} = \frac{(I-1)QMTat.}{\sigma^2} \;\sim\; \chi^2_{(I-1)}, \quad \text{sob} \quad H_0^{(\tau)} : \sigma^2_\tau = 0;$$

$$\frac{SQBlocos}{\sigma^2} = \frac{(J-1)QMBlocos}{\sigma^2} \;\sim\; \chi^2_{(J-1)}, \quad \text{sob} \quad H_0^{(\beta)} : \sigma^2_\beta = 0;$$

$$\frac{SQRes.}{\sigma^2} = \frac{(I-1)(J-1)QMRes.}{\sigma^2} \;\sim\; \chi^2_{[(I-1)(J-1)]}, \text{ independentemente de} H_0^{(\tau)} \text{ e } H_0^{(\beta)}.$$

4. De acordo com os resultado encontrados para as esperanças dos quadrados médios, na seção 2.6, observa-se que:

   - Sob $H_0^{(\tau)} : \sigma^2_\tau = 0$, o $QMTrat.$ é um estimador não viciado para σ^2 e que independentemente de que $H_0^{(\tau)}$ se verifique o QMRes também é um estimador não viciado para σ^2. Então pode-se deduzir que uma estatística apropriada para contrastar as hipóteses $H_0^{(\tau)} : \sigma_\tau^2 = 0$ *vs* $H_1^{(\tau)} : \sigma_\tau^2 > 0$ é aquela que compara o $QMTrat.$ com $QMRes.$, ou seja,

$$F_t = \frac{\frac{\frac{SQTrat.}{\sigma^2}}{(I-1)}}{\frac{\frac{SQRes.}{\sigma^2}}{(I-1)(J-1)}} = \frac{\frac{\frac{(I-1)QMTrat.}{\sigma^2}}{(I-1)}}{\frac{\frac{(I-1)(J-1)QMRes.}{\sigma^2}}{(I-1)(J-1)}} = \frac{QMTrat.}{QMRes.} \sim F_{[(I-1);\ (I-1)(J-1)]},$$

ou seja, a estatística F_t segue uma distribuição F com $(I-1)$ graus de liberdade do numerador e $(I-1)(J-1)$ graus de liberdade do denominador.

Dessa forma, se $H_0^{(\tau)} : \sigma_\tau^2 = 0$ for verdadeira, tanto o numerador quanto o denominador de F_t são estimadores não viciado para σ^2, no entanto, se $H_0^{(\tau)} : \sigma_\tau^2 = 0$ for falsa, a esperança matemática do $QMTrat.$ será maior que σ^2. Assim sendo, rejeita-se $H_0^{(\tau)} : \sigma_\tau^2 = 0$, quando o valor de F_t for maior que o correspondente valor teórico da distribuição F com $(I-1)$ e $(I-1)(J-1)$ grau de liberdade e nível de significância α, a qual será denotado por: $F_{[(I-1);\ (I-1)(J-1);\ \alpha]}$.

Portanto, a região de rejeição de $H_0^{(\tau)}$ será dada por:

$$RC = \{F_t \in \mathbb{R}^+ \,|\, F_t > F_{[(I-1);\ (I-1)(J-1);\ \alpha]}\}.$$

- De modo semelhante, tem-se que, sob $H_0^{(\beta)} : \sigma_\beta^2 = 0$, $QMBlocos$ é um estimador não viciado para σ^2, enquado que independentemente de que $H_0^{(\beta)}$ se verifique, $QMRes.$ é, também, um estimador não viciado para σ^2. Assim sendo, a estatística de teste para contrastar as hipóteses $H_0^{(\beta)} : \sigma_\beta^2 = 0$ *vs* $H_0^{(\beta)} : \sigma_\beta^2 > 0$ será:

$$F_b = \frac{\frac{\frac{SQBocos}{\sigma^2}}{(J-1)}}{\frac{\frac{SQRes.}{\sigma^2}}{(I-1)(J-1)}} = \frac{\frac{\frac{(J-1)QMBlocos}{\sigma^2}}{(J-1)}}{\frac{\frac{(I-1)(J-1)QMRes.}{\sigma^2}}{(I-1)(J-1)}} = \frac{QMBlocos}{QMRes.} \sim F_{[(J-1);\ (I-1)(j-1)]},$$

isto é, a estatística F_b segue uma distribuição F com $(J-1)$ graus de liberdade do numerador e $(I-1)(J-1)$ graus de liberdade do denominador.

De modo análogo, se $H_0^{(\beta)} : \sigma_\beta^2 = 0$ for verdadeira, tanto o numerador quanto o denominador de F_b são estimadores não viciado para σ^2, no entanto, se $H_0^{(\beta} : \sigma_\beta^2 = 0$ for falsa, a esperança matemática do $QMBlocos$ será maior que σ^2. Portanto, rejeita-se $H_0^{(\beta)} : \sigma_\beta^2 = 0$, quando o valor de F_b for maior que o correspondente valor teórico da distribuição F com $(J-1)$ e $(I-1)(J-1)$ grau de liberdade e nível de significância α, a qual será denotado por: $F_{[(J-1);\ (I-1)(J-1);\ \alpha]}$. Sendo assim, a região de rejeição de $H_0^{(\beta)}$ será dada por:

$$RC = \{F_b \in \mathbb{R}^+ \,|\, F_b > F_{[(J-1);\,(I-1)(J-1);\,\alpha]}\}.$$

2.8 Estimação por ponto dos componentes da variância

Conforme Montgomery, num modelo de efeitos aleatórios, usualmente é necessário estimar os componentes da variância. Neste sentido, tem-se que a variância de qualquer observação experimental, y_{ij}, é chamada de variância total e é denotada por σ_T^2, é dada por:

$$\sigma_T^2 = Var(y_{ij}) = \sigma_\tau^2 + \sigma_\beta^2 + \sigma^2. \tag{2.40}$$

O procedimento para estimar σ_τ^2, σ_β^2 e σ^2 é chamado de *método da análise da variância*, que consiste em utilizar as linhas da tabela da análise da variância. Tal método não exige a hipótese de normalidade das observações. O procedimento consiste em igualar os valores esperados dos quadrados médios a seus valores observados na análise da variância e resolver a equação em relação aos componente da variância. Para este caso, tem-se que:

$$\begin{aligned} V_1 &= QMTrat. &&= \sigma^2 + J\sigma_\tau^2 \\ V_2 &= QMBlocos &&= \sigma^2 + I\sigma_\beta^2 \\ V_3 &= QMRes. &&= \sigma^2. \end{aligned}$$

Assim, os estimadores dos componentes da variância são:

$$\begin{aligned} \hat{\sigma}^2 &= V_3 &&= QMRes. \\ \hat{\sigma}_\tau^2 &= \tfrac{V_1 - V_3}{J} &&= \tfrac{QMTrat - QMRes}{J} \\ \hat{\sigma}_\beta^2 &= \tfrac{V_2 - V_3}{I} &&= \tfrac{QMBlocos - QMRes.)}{I}. \end{aligned} \tag{2.41}$$

Portanto, o estimador da variabilidade total é dado por:

$$\hat{\sigma}_T^2 = \hat{\sigma}_\tau^2 + \hat{\sigma}_\beta^2 + \hat{\sigma}^2. \tag{2.42}$$

2.8.1 Estimadores da proporção da variabilidade total explicada pelos componente da variância

A expressão (2.42) sugere que, se todas as estimativas dos componentes da variância são positivas, o estimador da proporção da variância total explicada pelos tratamentos e pelos blocos, seja:

$$R^2 = \frac{100(\hat{\sigma}^2_\tau + \hat{\sigma}^2_\beta)}{\hat{\sigma}^2_T} \tag{2.43}$$

e para cada componente de variância individual, seja:

$$\begin{cases} R^2_\tau = \frac{100\hat{\sigma}^2_\tau}{\hat{\sigma}^2_T} \longrightarrow \text{estimador da proporção da variação total explicada pelos tratamentos,} \\ \\ R^2_\beta = \frac{100\hat{\sigma}^2_\beta}{\hat{\sigma}^2_T} \longrightarrow \text{estimador da proporção da variação total explicada pelos blocos,} \\ \\ R^2_\epsilon = \frac{100\hat{\sigma}^2}{\hat{\sigma}^2_T} \longrightarrow \text{estimador da proporção da variação total explicada pelo acaso.} \end{cases} \tag{2.44}$$

Observação: Note que, quando há uma ou mais estimativas de componentes de variância negativas, não é simples interpretar os resultados fornecidos pelas expressões dadas em (2.44).

2.8.2 Estimativa negativa de um componente da variância

algumas vezes, o método da análise de variância produz uma estimativa negativa de um componente da variância. Como os componentes da variância são, por definição, não-negativos, uma estimativa negativa de um componente da variância é inaceitável. Uma alternativa é aceitar a estimativa e usá-la como evidência de que o verdadeiro valor do componente da variância é zero, adimitindo-se que a variação amostral levou a uma estimativa negativa. Embora isso tenha um apelo intuitivo, atrapalhará as propriedades estatísticas de outras estimativas. Outra alternativa é reestimar o componente negativo da variância por um método que sempre resulte em estimativas não-negativas. Outra possibilidade é ainda considerar a estimativa negativa como evidência de que o modelo linear proposto é incorreto, o que exige um estudo do modelo e de suas hipóteses para a determinação de um modelo mais apropriado.

2.9 Estimação por intervalo dos componentes da variância

De acordo com Barbin, Considerando-se os elementos da **Tabela** 3, os intervalos com $(1-\alpha)100\%$ de confiança para os componentes de variância são, obtidos como segue:

$$IC(\sigma^2)_{[(1-\alpha)100\%]} : \left[\frac{\nu_e\hat{\sigma}^2}{\chi^2_{[\nu_e;\,(1-\alpha/2)]}}\ ;\ \frac{\nu_e\hat{\sigma}^2}{\chi^2_{[\nu_e;\,\alpha/2]}}\right],\ \text{onde, } \nu_e \text{ é o número de GL dos resíduos;}$$

$$IC(\sigma^2_\tau)_{[(1-\alpha)100\%]} : \left[\frac{\nu'_t\hat{\sigma}^2_\tau}{\chi^2_{[\nu'_t;\,(1-\alpha/2)]}}\ ;\ \frac{\nu'_t\hat{\sigma}^2_\tau}{\chi^2_{[\nu'_t;\,\alpha/2]}}\right];\ \text{onde, } \nu'_t \text{ é obtido por: } \nu'_t = \frac{[QMTrat.-QMRes.]^2}{\frac{(QMTrat.)^2}{(I-1)}+\frac{(QMRes.)^2}{(I-1)(J-1)}} \tag{2.45}$$

$$IC(\sigma^2_\beta)_{[(1-\alpha)100\%]} : \left[\frac{\nu'_b\hat{\sigma}^2_\beta}{\chi^2_{[\nu'_b;\,(1-\alpha/2)]}}\ ;\ \frac{\nu'_b\hat{\sigma}^2_\beta}{\chi^2_{[\nu'_b;\,\alpha/2]}}\right],\ \text{onde, } \nu'_b \text{ é obtido por: } \nu'_b = \frac{[QMBlocos-QMRes.]^2}{\frac{(QMBlocos)^2}{J-1}+\frac{(QMRes.)^2}{(I-1)(J-1)}}.$$

As expressões usadas para calcular os graus de liberdade, ν'_t e ν'_b, associados às respectivas estimativas dos componentes de variâncias, $\hat{\sigma}^2_\tau$ e $\hat{\sigma}^2_\beta$ em (2.44), são conhecidas por *fórmulas de Satterthwaite.*[1]

2.10 Procedimento prático para a análise dos dados de um experimento em blocos ao acaso de efeitos aleatórios

Um procedimento prático a ser utilizado na análise dos dados de um experimento em blocos ao acaso com efeitos aleatórios, pode ser o seguinte:

1. Enunciar as hipóteses:
$$H_0^{(\tau)} : \sigma^2_\tau = 0 \quad vs \quad H_1^{(\tau)} : \sigma^2_\tau > 0$$
e
$$H_0^{(\beta)} : \sigma^2_\beta = 0 \quad vs \quad H_1^{(\beta)} : \sigma^2_\beta > 0.$$
2. Fixar o nível de significância α e observar as estatísticas dos testes, que são
$$F_t = \frac{QMTrat.}{QMRes.} \ \sim\ F_{[(I-1);\,(I-1)(J-1);\,\alpha]}$$
e
$$F_b = \frac{QMBlocos}{QMRes.} \ \sim\ F_{[(J-1);\,(I-1)(J-1);\,\alpha]}.$$
3. Com o auxilio da tabela da distribuição de probabilidade da F estabelecer as regiões de rejeição das hipóteses $H_0^{(\tau)}$ e $H_0^{(\beta)}$. Isto é,

[1] Ver Barbin(1993), pg. 44-48.

$$RC_t = \{F_t \in \mathbb{R}^+ | F_t \geq F_{[(I-1);\ (I-1)(J-1);\ \alpha]}\}$$

e

$$RC_b = \{F_b \in \mathbb{R}^+ | F_b \geq F_{[(J-1);\ (I-1)(J-1);\ \alpha]}\}.$$

4. Com os dados de um experimentais real, organizar a tabela da análise da variância, para obter os valores das variáveis de teste, ou seja,

F. Variação	GL	SQ	QM	F
Tratamentos	$I-1$	$SQTrat.$	$QMTrat.$	$F_t = QMTrat./QMRes.$
Blocos	$J-1$	$SQBlocos$	$QMBlocos$	$F_b = QMBlocos/QMRes.$
Resíduo	$(I-1)(J-1)$	$SQRes.$	$QMRes.$	-
Total	$IJ-1$	$SQTotal$	-	-

onde,

$$SQTotal = \sum_{i,j} y_{ij}^2 - C, \quad \text{e} \quad C = \frac{\left(\sum_{i,j} y_{ij}\right)^2}{IJ};$$

$$SQTrat. = \frac{1}{J}\sum_{i} y_{i.}^2 - C \quad \text{e} \quad QMTrat. = SQTrat./(I-1);$$

$$SQBlocos = \frac{1}{I}\sum_{j} y_{.j}^2 - C \quad \text{e} \quad QMBlocos = SQBlocos/(J-1);$$

$$SQRes. = SQTotal - SQTrat. - SQRes. \quad \text{e} \quad QMRes. = SQRes./(I-1)(J-1).$$

Concluir: Se $F_t > F_{[(I-1);\ (I-1)(J-1);\ \alpha]}$ rejeita-se $H_0^{(\tau)} : \sigma_\tau^2 = 0$ e conclui-se, ao nível de significância α, que as variações na variável resposta devidas aos efeitos dos tratamentos são estatisticamente significativas (isto é, aceita-se $H_0^{(\tau)} : \sigma_\tau^2 > 0$). De modo análogo, se $F_b > F_{[(J-1);\ (I-1)(J-1)]}$, rejeita-se $H_0^{(\beta)} : \sigma_\beta^2 = 0$ e conclui-se, ao nível de significância α, que as variações na variável resposta provocadas pelos efeitos dos blocos são estatísticamente significativas (ou seja, aceita-se $H_0^{(\beta)} : \sigma_\beta^2 > 0$).

5. Estimar os componentes da variância;

6. Estimar as proporções da variabilidade total explicadas pelos componentes da variância individualmente;

7. Construir os intervalos com $(1-\alpha)\times 100\%$ de confiça para cada componente da variância;

8. Fazer comentários sobre os achados na pesquisa.

2.11 Uma aplicação da teoria a uma pesquisa real

Na **Tabela** 4 encontram-se as produções, em toneladas por hectare, de 10 cultivares de arroz ($I = 10$), considerados num experimento em blocos casualizados, com três blocos ($J = 3$). Adaptado de Magno.

Tabela 4: Dados da produão de grãos (ton/ha) de 10 cultivares de arroz obtidos de um experimento em blocos casualizados com tês repetições, Lavras-MG.

Cultivares (C_i)	**Blocos** B_1	B_2	B_3	**Totais** ($y_{i.}$)	**Médias** ($\bar{y}_{i.}$)
C_1	5,43	5,98	5,40	16,81	5,60
C_2	3,72	3,58	4,77	12,07	4,02
C_3	4,37	3,79	3,75	11,91	3,97
C_4	4,59	5,31	5,76	15,66	5,22
C_5	4,74	4,00	4,81	13,55	4,52
C_6	3,23	2,93	3,00	9,16	3,05
C_7	4,32	4,17	4,79	13,28	4,43
C_8	4,38	3,74	3,65	11,77	3,92
C_9	6,37	5,08	5,29	16,74	5,58
C_{10}	4,84	4,20	4,75	13,79	4,60
Totais	45,99	42,78	45,97	134,74	4,49

2.11.1 Cálculos das somas dos quadrados e a análise da variância

A partir dos dados experimentais calcula-se:

$$C = \frac{\left(\sum_{i,j} y_{ij}\right)^2}{IJ} = \frac{(134{,}74)^2}{10\times 3} = 605,1623;$$

$$SQTotal = \sum_{i,j} y_{ij}^2 - C = 626,8852 - 605,1623 = 21,7229;$$

$$SQTrat. = \frac{1}{J}\sum_i y_{i.}^2 - C = \frac{1}{3}\left[(16,81)^2 + \cdots + (13,79)^2\right] - 605,1623 = 17,5497;$$

$$SQBlocos = \frac{1}{I}\sum_j y_{.j}^2 - C = \frac{1}{10}\left[(45,99)^2 + (42,78)^2 + (45,97)^2\right] - 605,1623 = 0,6827;$$

$$SQRes. = SQTotal - SQTrat. - SQRes. = 21,7229 - 17,5497 = 3,4905.$$

e em seguida, organiza-se os resultados na tabela da análise da variância, **Tabela** 5.

Tabela 5: A análise da variância

F. Variação	GL	SQ	QM	F
Tratamentos	9	17,5497	1,9500	$F_t = 9,53$
Blocos	2	0,6827	0,3413	$F_b = 1,67$
Resíduos	18	3,6827	0,2046	-
Total	29	21,7229	-	-

Conclusões: Conforme pode ser observado, Os resultados apresentados na **Tabela** 5, levam às seguintes conclusões: Como $F_t = 9,53 > F_{[9;\,18;\,0,01]} = 3,71$, então rejeita-se $H_0^{(\tau)} : \sigma_\tau^2 = 0$ e conclui-se, ao nível de significância de 1%, que as variações na variável resposta provocadas pelos efeitos dos tratamentos são estatisticamente significativas (isto é, aceita-se $H_1^{(\tau)} : \sigma_\tau^2 > 0$). De modo análogo, como $F_b = 1,67 < F_{[2;\,18;\,0,05]} = 3,55$, então aceita-se $H_0^{(\beta)} : \sigma_\beta^2 = 0$ e conclui-se, ao nível de significância de 5%, que as variações na variável resposta devidas aos efeitos dos blocos não são estatísticamente significativas.

2.11.2 Estimativas dos componentes da variância

De acordo com as equações (2.41) e (2.42), as estimativas dos componentes da variância encontradas foram:

$$\hat{\sigma}^2 = QMRes. = 0,2046;$$

$$\hat{\sigma}_\tau^2 = \frac{QMTrat. - QMRes.}{J} = \frac{1,9500 - 0,2046}{3} = 0,5818;$$

$$\hat{\sigma}_\beta^2 = \frac{QMBlocos - QMRes.}{I} = \frac{0,3413 - 0,2046}{10} = 0,0137;$$

$$\hat{\sigma}_T^2 = \hat{\sigma}_\tau^2 + \hat{\sigma}_\beta^2 + \hat{\sigma}^2 = 0,5818 + 0,0137 + 0,2046 = 0,8001.$$

Conclusões: Conforme pode ser observado, estima-se que a variabilidade provocada na produção de arroz, devidas aos efeitos dos cultivares ($\hat{\sigma}_\tau^2 = 0,5818$) é quase *três* vezes maior que a variabilidade provocada pelos efeitos do acaso ($\hat{\sigma}^2 = 0,2046$) e quase 43 vezes maior que a variabilidade provocada pelos efeitos dos blocos ($\hat{\sigma}_\beta^2 = 0,0137$).

2.11.3 Estimativas das proporções da variabilidade total explicada pelos componentes da variância

A partir dos resultados encontrados na seção 2.11.2, é possível estimar as prporções da variação total explicada pelos componentes individualmente, a saber:

$$R^2 = \frac{100(\hat{\sigma}^2_\tau + \hat{\sigma}^2_\beta)}{\hat{\sigma}^2_T} = \frac{100(0{,}5818+0{,}0137)}{0{,}8001} = 74,4\%;$$

$$R^2_\tau = \frac{100\hat{\sigma}^2_\tau}{\hat{\sigma}^2_T} = \frac{100\times 0{,}5818}{0{,}8001} = 72,7\%;$$

$$R^2_\beta = \frac{100\hat{\sigma}^2_\beta}{\hat{\sigma}^2_T} = \frac{100\times 0{,}0137}{0{,}8001} = 1,7\%;$$

$$R^2_\epsilon = \frac{100\hat{\sigma}^2}{\hat{\sigma}^2_T} = \frac{100\times 0{,}2046}{0{,}8001} = 25,6\%.$$

Conclusões: Estima-se que a variabilidade total observada na produção do arroz, 72,7% são devidas aos efeitos dos cultivares, 1,7% são provocadas pelos efeitos dos blocos e 25,6% pelo efeito do acaso.

2.11.4 Estimativas dos intervalos com 95% de confiança para os componentes da variância

Usando os resulados da **Tabela** 5 e as expressões dadas em (2.45), tem-se que:

$$IC(\sigma^2)_{[95\%]} : \left[\frac{18\times 0,2046}{\chi^2_{[18;\,0{,}975]}}\ ;\ \frac{18\times 0,2046}{\chi^2_{[18;\,0{,}025]}}\right] = \left[\frac{3,6828}{31,50}\ ;\ \frac{3,6828}{8,23}\right] = [0,1169;\ 0,4475];$$

$$IC(\sigma^2_\tau)_{[95\%]} : \left[\frac{8\times 0,5818}{\chi^2_{[8;\,0{,}975]}}\ ;\ \frac{8\times 0,5818}{\chi^2_{[8;\,0{,}025]}}\right] = \left[\frac{4,6544}{17,50}\ ;\ \frac{4,6544}{2,18}\right] = [0,2660;\ 2,1350],$$

uma vez que,

$$\nu'_t = \frac{[QMTrat. - QMRes.]^2}{\frac{(QMTrat.)^2}{(I-1)} + \frac{(QMRes.)^2}{(I-1)(J-1)}} = \frac{[1,9500 - 0,2046]^2}{\frac{(1{,}9500)^2}{9} + \frac{(0{,}2046)^2}{18}} = 7,1710 \approx 8.$$

$$IC(\sigma^2_\beta)_{[95\%]} : \left[\frac{1\times 0,0137}{\chi^2_{[1;\,0{,}975]}}\ ;\ \frac{1\times 0,0137}{\chi^2_{[1;\,0{,}025]}}\right] = \left[\frac{0,0137}{5,02}\ ;\ \frac{0,0137}{0,001}\right] = [0,0027;\ 13,7000],$$

tendo em vista que,

$$\nu'_b = \frac{[QMBlocos - QMRes.]^2}{\frac{(QMBlocos)^2}{J-1} + \frac{(QMRes.)^2}{(I-1)(J-1)}} = \frac{[0,3413 - 0,2046]^2}{\frac{(0{,}3413)^2}{2} + \frac{(0{,}2046)^2}{18}} = 0,3085 \approx 1.$$

Interpretação:

Se forem instalados um grande número de experimentos em condições identênticas ao ensaio aqui analizado e para cada um deles construírem-se os intervalos do tipo, $\left[\frac{\nu_e\hat{\sigma}^2}{\chi^2_{[\nu_e;\,(1-\alpha/2)]}}\ ;\ \frac{\nu_e\hat{\sigma}^2}{\chi^2_{[\nu_e;\,\alpha/2]}}\right]$, $\left[\frac{\nu_t'\hat{\sigma}^2_\tau}{\chi^2_{[\nu_t';\,(1-\alpha/2)]}}\ ;\ \frac{\nu_t'\hat{\sigma}^2_\tau}{\chi^2_{[\nu_t';\,\alpha/2]}}\right]$ e $\left[\frac{\nu_b'\hat{\sigma}^2_\beta}{\chi^2_{[\nu_b';\,(1-\alpha/2)]}}\ ;\ \frac{\nu_b'\hat{\sigma}^2_\beta}{\chi^2_{[\nu_b';\,\alpha/2]}}\right]$, então em $(1-\alpha)\times 100\%$ deles conterão os verdadeiros valores dos componentes da variância σ^2, σ^2_τ e σ^2_β, respectivamente. Assim, para esse experimento, pode-se concluir que há 95% de confiança de que o intervalo $[0,1169;\ \ 0,4475]$ contenha o verdadeiro valor de σ^2, 95% de confiança de que o intervalo $[0,2660;\ \ 2,1350]$ contenha o verdadeiro valor de σ^2_τ e 95% de confiança de que o intervalo $[0,0027;\ \ 13,7000]$ contenha o verdadeiro valor de σ^2_β.

3 Conclusões

Ao desenvolver a teoria estatística para as análises dos dados de um experimentos em blocos completos casualizados de efeitos aleatórios, evidenciou-se claramente os seguintes aspectos:

1. O modelo matemático associado aos dados do experiemnto onde os tratamentos e os blocos são de efeitos aleatórios é exatamente o mesmo, daqueles nos quais se considera que os efeitos dos tratamentos ou blocos ou ambos, sobre a variável resposta, são de efeitos fixos. Diferenciando-se apenas, no que se refere às suposições acerca dos termos no modelo;

2. O método utilizado para obtenção dos estimadores dos efeitos dos tratamentos e dos blocos, foi o método da máxima verossimilhança, tendo em vista que os estimadores obtidos a partir deste método, apresentam excelentes propriedades e, além disso, sugere de pronto a estimação da variância do erro, σ^2, enquanto que o método dos mínimos quadrados não deixa claro que é necessário estimá-lo;

3. O que se pretende inferir sobre os efeitos dos tratamentos e/ou blocos numa pesquisa baseada num experimento de efeitos aleatórios é diferente daquilo que se pretende com um experimento onde os efeitos são considerados de efeitos fixos. No primeniro caso, o que se pretende, é verificar se os efeitos dos tratamentos ou dos blocos, sobre a variável resposta, não variam de um tratamento para outro ou de um bloco para outro. Em linguagem estatística, estas ideias serão formalizadas, através da contrastação das seguintes hipóteses: $H_0^{(\tau)} : \sigma_\tau^2 = 0 \; vs \; H_1^{(\tau)} : \sigma_\tau^2 > 0$ e $H_0^{(\beta)} : \sigma_\beta^2 = 0 \; vs \; H_1^{(\beta)} : \sigma_\beta^2 > 0$. De acordo com a literatura, no caso onde os efeitos dos tratamentos e blocos são considerados fixos, o que se pretende é verificar se os efeitos dos tratamentos ou dos blocos, sobre a variável resposta, não existem. Formalmente, seria contrastar as seguintes hipóteses: $H_0^{(\tau)} : \tau_1 = \tau_2 = \cdots = \tau_I = 0 \; vs \; H_1^{(\tau)} : \tau_i \neq 0$, para pelo menos um τ_i, $i = 1, \cdots, I$ que é equivalente a contrastar as hipóteses $H_0^{(\tau)} : \mu_1 = \mu_2 = \cdots = \mu_I = \mu \; vs \; H_1^{(\tau)} : \mu_i \neq \mu_{i'}$, para pelo menos um par de médias $(\mu_i,\ \mu_{i'})$, $i \neq i';\ i, i' = 1, \cdots, I$.

4. Após a constatação da existência da variabilidade dos efeitos dos tratamentos e blocos, sobre a variável resposta, isto é, após a rejeição das hipóteses, $H_0^{(\tau)}$ e $H_0^{(\beta)}$, a um nível de significância α desejado, as análises estatísticas complementares, serão baseadas em estimação dos componentes da variância, coeficientes de explicação e construção de intervalos de confiaça para os componentes de variância. Por outro lado, de acordo com a literatura, se os efeitos dos tratamentos e blocos forem considerados fixos, as análises estatísticas serão complementadas aplicando-se um teste de comparações múltipas das médias, por exemplo: *Tukey, t-Student, Duncan, Scheffé*, dentre outros;

5. É Oportuno enfatizar a importância da validação das suposições impostas aos termos no modelo matemático utilizado para descrever as observações experimentais. As conclusões acerca dos achados na pesquisa só deverão ser consideradas verdadeiras após a comprovação estatística das suposições: *Aditividade, Normalidade, Homocedasticidade, Independência dos erros*.

6. As análises estatísticas dos dados do experimento utilizado para ilustrar a teoria, constataram que a variabilidade dos efeitos dos tratamentos é estatisticamente significativa, ao nível de 1% de significância, enquanto que a variabilidade dos efeitos dos blocos não apresentaram significância estatística, ao nível de 5% de probabilidade pelo teste F. Ao empregar o método da ANOVA, encontrou-se as seguintes estimativas para os componentes da variância: $\hat{\sigma}_\tau^2 = 0,5818$, $\hat{\sigma}_\beta^2 = 0,0137$ e $\hat{\sigma}^2 = 0,2046$ estes valores estimam que 72,7% da variância total é explicada pelos efeitos dos tratamentos, 1,7% é explicada pelos efeitos dos blocos e 25,6% pela variação do acaso. A estimativa da variância total foi $\hat{\sigma}_T^2 = \hat{\sigma}_\tau^2 + \hat{\sigma}_\beta^2 + \hat{\sigma}^2 = 0,8001$.

7. Há dificuldade de se encontrar na literatura especializada, textos que abordem de modo completo, a teoria estatística que foi proposta e desenvolvida nesta monografia. Embora os resultados encontrados estejam em concordância com aqueles disponíveis na literatura, é possível que alguns equívocos, do ponto de vista teórico, tenham sido cometidos.

Referências

BARBIN, D. *Componentes de Variância (Teoria e Aplicações)*. Ed. FEALQ (Fundação de Estudos Agrários Luiz de Queiroz). 2ed. Piracicaba - São Paulo, 1988, 86p.

GOMEZ, K.A.; GOMEZ, A.A. *Statístical produceres for Agricultural Research*. Ed. John Wiley, New York, 1984, 680p.

HOGG, R.V.; CRAIG,A.A. *Introduction to Mathematical Statistics.* . 4ed. New York: Macmillan Publishing Co. Inc., 1989. 438 p.

LEAL, J.G.; PORRAS, A.M.L. *Deseño estadístico de experimentos - Análises de la varianza.* Granada - ES: Grupo Editorial Universitário. 1998. 357 p.

MONTGOMERY, D.C. *Design and Analysis of Experiments.* Singapore: Post & Telecon Press, 2007. 642 P.

MOOD, A.M.; GRAYBILL, F.A.; BOES, D.C. *Introduction to the theory of statistics*. 3.Ed. Singapore: Mcgraw-Hill, 1974. 463 p.(Statistical Series)

MOORE, D.S.; MCCABE, G.P. *Introdução à prática da Estatística* . 3. Ed. Rio de Janeiro LTC, 2002. 536 p.

RAMALHO, M.A.P.; FERREIRA, D.F.; OLIVEIRA, A.C. *Experimentação em Genética e melhoramentos de plantas*. 2. Ed. Lavras - MG: UFLA, 2005. 300 p.

ROHATGI,V.K. *An Introduction to probability Theory and Mathematical Statistics.* New York: John Wiley, 1976. 684 p.

SANTOS, J.W.; ALMEIDA, F.A.C.; BELTRÃO, N.E.M.; CAVALCANTI, F.B. *Estatística Experimental Aplicada*. 2ed. UFCG, EMBRAPA. Campina Grande - Campina Grande, 2008.

VANDERLEI, P.F. *Estatística Experimental aplicada à agronomia*. 3.Ed. edUFAL, Maceió, 2000, 419 p.