



Universidade Estadual da Paraíba
Centro de Ciências e Tecnologia
Departamento de Estatística

Bruno Henrique Gomes dos Santos

Aspectos teóricos e práticos com aplicação da análise
estatística de um experimento em blocos completos
casualizados com repetições dentro dos blocos

Campina Grande
Dezembro de 2012.

Bruno Henrique Gomes dos Santos

**Aspectos teóricos e práticos com aplicação
da análise estatística de um experimento em
blocos completos casualizados com
repetições dentro dos blocos**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento as exigências legais para obtenção do título de Bacharel em Estatística.

Orientador:
João Gil de Luna

Campina Grande
Dezembro de 2012.

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL – UEPB

S237a Santos, Bruno Henrique Gomes dos.
Aspectos teóricos e práticos com aplicação da análise estatística de um experimento em blocos completos casualizados com repetições dentro dos blocos [manuscrito] / Bruno Henrique Gomes dos Santos. – 2012.
63f. : il.

Trabalho de Conclusão de Curso (Graduação em Estatística) – Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2012.
“Orientação: Prof. Dr. João Gil de Luna, Departamento de Estatística”.

1. Estatística Experimental. 2. Probabilidade. 3. Pesquisa Experimental. I. Título.

21. ed. CDD 519.2

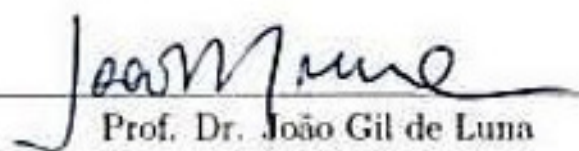
Bruno Henrique Gomes dos Santos

**Aspectos teóricos e práticos com aplicação
da análise estatística de um experimento em
blocos completos casualizados com
repetições dentro dos blocos**

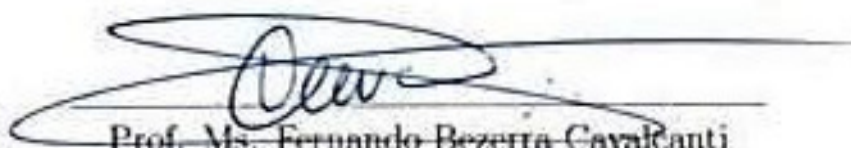
Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento as exigências legais para obtenção do título de Bacharel em Estatística.

Aprovado em: 10/12/2012

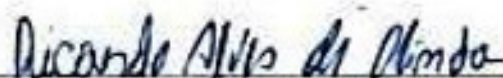
Banca Examinadora:



Prof. Dr. João Gil de Luna
Orientador
Universidade Estadual da Paraíba



Prof. Ms. Fernando Bezerra Cavalcanti
Universidade Estadual da Paraíba



Prof. Dr. Ricardo Alves de Olinda
Universidade Estadual da Paraíba

Dedicatória

Dedico este trabalho a minha esposa Cida, a meu querido filho Lucas e a minha enteadada Mariana que me impulsionaram a buscar vida nova a cada dia, concedendo a mim a oportunidade de me realizar ainda mais. Dedico também a minha mãe Cristina, que em nenhum momento mediu esforços para realização dos meus sonhos, que me guiou pelos caminhos corretos, me ensinou a fazer as melhores escolhas, me mostrou que a honestidade e o respeito são essenciais à vida, e que devemos sempre lutar pelo que queremos.

Obrigado!

Agradecimentos

Agradeço primeiramente a Deus por ter me dado forças para enfrentar os obstáculos que insistiam em aparecer no decorrer dessa caminhada, fazendo-me concluir um grande passo em minha vida.

Agradeço a minha família por terem aceito se privar de minha companhia pelos estudos, compreendendo que buscava a todo tempo o melhor para todos.

A UEPB por todos os recursos oferecidos durante o curso, conseguindo assim êxito em todas as pesquisas.

Ao Prof. Gil pela imensa atenção e dedicação ao nosso trabalho, e o encorajamento nos momentos difíceis, como também a todo conhecimento passado por ele em diversos campos da estatística. Ao Prof. Gustavo por todo aprendizado proporcionado por ele me fazendo ter sempre confiança em trabalhos realizados extra classe.

A todos os meus colegas de classe que juntos transferiram conhecimentos para que fosse possível a conclusão do nosso curso. Que essas amizades durem tanto quanto foram intensas.

Resumo

A pesquisa experimental é amplamente utilizada em diversas áreas do conhecimento, para tal é desenvolvido um método em que o pesquisador intervém na amostra, impondo deliberadamente os níveis de uma ou mais características explanatórias com o propósito de encontrar inferências referentes aos efeitos causais dessas características sobre características respostas. Essas características explanatórias são denominadas características de tratamento e seus níveis, tratamentos. Exemplos comuns de tratamentos são diferentes estímulos apresentados ou impostos a animais ou plantas, tais como diferentes dietas administradas a animais ou diferentes fungicidas aplicados a plantas. As conclusões desses experimentos são obtidas utilizando-se da estatística experimental, estatística essa que usa os dados coletados para inferir resultados com o objetivo de aprimorar ou até mesmo, quando necessário, refazer o experimento. Neste trabalho aborda-se todo o desenvolvimento teórico dos procedimentos que dão suporte a uma análise estatística dos dados de um experimento em blocos completos casualizados de efeitos fixos e com repetições dentro dos blocos. Será apresentado um possível desenho desse tipo de experimento no campo, juntamente com a tabela para o recolhimento dos dados, defini-se o modelo matemático para descrever as observações experimentais, utiliza-se o método de mínimos quadrados para encontrar os estimadores dos termos do modelo, apresenta-se os resultados da decomposição da variabilidade total que são organizados na tabela da análise da variância (ANOVA), estuda-se as distribuições de probabilidade dos estimadores e por fim calcula-se os valores esperados das somas de quadrados. Por fim, um exemplo real será utilizado para ilustrar a metodologia, e os resultados serão discutidos e interpretados convenientemente.

Palavras-chave: Estatística Experimental, Soma de Quadrados, ANOVA.

Abstract

The experimental research is widely used in various areas of knowledge, such a method is developed in which the researcher intervenes in the sample, levels of deliberately imposing one or more characteristics explanatory in order to find causal inferences regarding the effects of these characteristics on response characteristics. These features are termed features explanatory treatment and its levels treatments. Examples of common treatments are different stimuli or imposed animals or plants, such as different diets administered to animals or applied to plant fungicides. The conclusions of these experiments are obtained using the experimental statistics, this statistic that uses the collected data to infer results in order to enhance or even, when necessary, redo the experiment. This paper addresses to the entire theoretical development of procedures that support a statistical analysis of an experiment in randomized complete block design with fixed effects and replicates within the blocks. We will present a possible design of this type of experiment in the field, along with the table for the collection of data, set up the mathematical model to describe the experimental observations, we use the method of least squares estimators to find the terms of the model presents the results of the decomposition of the total variance that are arranged in the table analysis of variance (ANOVA) is studied probability distributions of the estimators and finally calculates the expected values of the sums of squares. Finally, a real exmplo will be used to illustrate the methodology and results discussed and interpreted properly.

Keywords: Experimental Statistics, Sum of Squares, ANOVA.

Sumário

1	Inrodução	p. 9
2	Fundamentação Teórica	p. 12
2.1	Um possível desenho do experimento no campo	p. 12
2.2	Organização dos dados experimentais	p. 13
2.3	O Modelo matemático	p. 13
2.4	Estimação dos parâmetros, dos erros e das observações	p. 14
2.5	Decomposição da variabilidade total	p. 17
2.6	Distribuição de probabilidade dos estimadores	p. 21
2.6.1	Distribuição de probabilidade de \bar{y}_{\dots} , o estimador de μ	p. 22
2.6.2	Distribuição de probabilidade da correção para média, C	p. 23
2.6.3	Distribuição de probabilidade de t_i , o estimador de τ_i	p. 23
2.6.4	Distribuição de probabilidade da SQ_{Trat}	p. 25
2.6.5	Distribuição de probabilidade de b_j , o estimador de β_j	p. 25
2.6.6	Distribuição de probabilidade da SQ_{Blocos}	p. 27
2.6.7	Distribuições de probabilidade de m_i , o estimador de $\mu_i = \mu + \tau_i$, (a média do tratamento i) e de m_j , o estimador de $\mu_j = \mu + \beta_j$, (a média do bloco j).	p. 27
2.6.8	Distribuições de probabilidade de contrastes de interesse	p. 28
2.6.9	Distribuição de probabilidade da soma de quadrados da h -ésima combinação linear das médias dos tratamentos	p. 30
2.6.10	Distribuição de probabilidade de $\hat{\epsilon}_{ij}$, o estimador de ϵ_{ij}	p. 31
2.6.11	Distribuição de probabilidade da $SQ_{Erro Entre}$	p. 36

2.6.12	Distribuição de probabilidade de $\hat{\varepsilon}_{ijr}$, o estimador de ε_{ijr}	p. 36
2.7	Valores Esperados das Somas de Quadrados	p. 38
2.8	Comparações múltiplas das médias duas a duas	p. 47
2.9	Análises estatísticas	p. 47
2.9.1	Hipóteses sobre tratamento	p. 48
2.9.2	Hipóteses sobre Bloco	p. 50
2.9.3	A tabela da ANOVA	p. 52
3	Aplicação da teoria a um exemplo real	p. 53
3.1	Descrição do conjunto de dados experimentais	p. 53
3.2	Cálculos das somas de quadrados e análise da variância	p. 54
3.3	Comprovação da idoneidade do modelo	p. 57
4	Conclusão Final	p. 60
	Referências	p. 62

1 Introdução

Tem-se registros que o método experimental remonta a pelo menos 4 séculos antes de Cristo, quando Aristóteles (384-322 a.C.) fez diversas descobertas referentes ao mundo natural, com base em experimentos, axiomas e argumentos filosóficos, ele concluiu, por exemplo, que a aceleração de um corpo em queda livre depende de sua massa, e que a terra devia ser uma esfera, já que a esfera é o sólido mais "perfeito". Porém foi no início do século XX com Ronald Aylmer Fisher (1890-1962), um jovem matemático do Colégio Caius de Cambridge, que iniciou-se o desenvolvimento do ramo da estatística relacionado com o planejamento e a análise de experimentos. Fisher lançou os fundamentos modernos da pesquisa experimental, as bases da inferência estatística e delineou muitos métodos originais para os vários problemas encontrados na Estação Experimental de Rothamsted, onde realizava seus trabalhos e em outras instituições de pesquisa. Introduziu diversas técnicas de análise de dados, como a análise da variação, que passou a ser amplamente utilizada na análise estatística de dados de experimentos, e a técnica de polinômios ortogonais para o uso de características ambientais.

A metodologia moderna da pesquisa experimental, desenvolvida a partir dos fundamentos e idéias lançados por Fisher para a pesquisa agrícola, teve muitos contribuintes em diversos países e passou a aplicar-se aos demais ramos da ciência e da tecnologia, tais como biologia, medicina, engenharia, indústria e ciências sociais.

Como conseqüência da origem da pesquisa experimental na agricultura, muito da terminologia ainda hoje utilizada compreende termos próprios da pesquisa agrícola. Assim, por exemplo, as designações "tratamento", "parcela" e "bloco" perderam suas conotações particulares da agricultura e são amplamente usadas na pesquisa experimental em muitas áreas da ciência (SILVA, 2007).

O delineamento em blocos ao acaso trata-se de um método para eliminar a heterogeneidade das unidades experimentais, e é o projeto mais fundamental em todos os tipos de experimentação. Historicamente, esse delineamento foi o primeiro projeto a estimar o erro

experimental e a testar a significância dos efeitos dos tratamentos, apesar da heterogeneidade das unidades experimentais em que as observações são adquiridas (LOVE, 1964). Os delineamentos experimentais são planejados de forma que a variação ao acaso seja reduzida o máximo possível. Os principais delineamentos são: *Inteiramente Casualizado*, *Blocos Completos Casualizados* e *Quadrados Latinos*.

Neste trabalho será abordado a teoria do *Delineamento em Blocos Completos Casualizados com repetições dentro dos Blocos*. O modelo matemático referente ao delineamento aqui estudado, propõe que os fatores sejam com interação e de efeito fixo.

Para obtenção dos estimadores dos efeitos envolvidos no modelo, utilizou-se o método da Mínimos Quadrados, esse método será escolhido porque é mais simples, e oferece os mesmos estimadores do de Máxima Verossimilhança.

Objetivos

Tem-se como principal objetivo desenvolver a teoria desse tipo de delineamento e tentar elucidar problemas eventualmente existentes nos experimentos com repetições dentro dos blocos, pois, essas teorias estatísticas que dão suporte as análises de dados de pesquisas experimentais são dificilmente encontradas na literatura, com isso a pouca aplicabilidade em pesquisas com um número razoavelmente grande de tratamentos, com tudo, o planejamento de experimento é de fundamental importância para a obtenção de resultados mais confiáveis além de proporcionar a diminuição da variabilidade e encontrar valores mais próximos dos esperados.

Para isso será utilizado um modelo matemático com o objetivo de representar e descrever o problema aqui colocado, de um experimento em blocos com repetições dentro dos blocos, estimando-se os termos do modelo com o uso do método de mínimos quadrados, que procura encontrar o melhor ajuste para um conjunto de dados tentando-se minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados, com os resultados encontrados iremos decompor a variabilidade total e por fim encontraremos as somas de quadrados, dispostas na tabela da Análise de Variância (ANOVA).

Com base em análises estatísticas, será mostrado que os estimadores do modelo matemático seguem todas as distribuições de probabilidade normal e que as somas de quadrados seguem todas as distribuições de probabilidade qui-quadrada, com essas definições expostas pode-se definir os valores esperados das somas de quadrado que irá ajudar

a definir as estatísticas de teste utilizadas na contrastação das hipóteses de interesse para tratamentos e para blocos, mostrando-se que o interesse maior em um experimento como esse é fazer inferências no efeito dos tratamentos, pois, será visto que os blocos por serem ambientes homogêneos não teram efeito sobre os tratamentos apresentados.

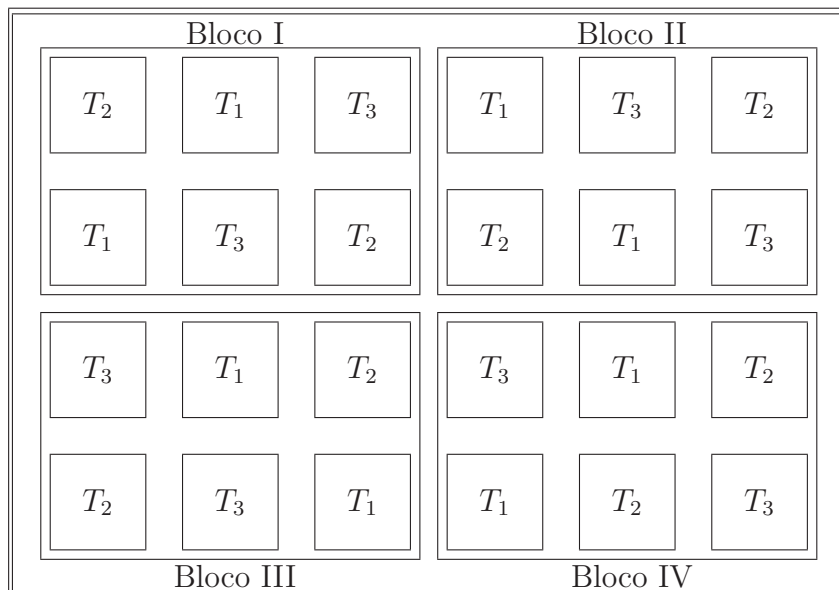
Será apresentado algumas estatísticas de teste que podem ser utilizadas para se fazer inferências marginais para alguns estimadores de tratamentos e de blocos.

2 Fundamentação Teórica

O foco principal deste trabalho é apresentar de modo claro o desenvolvimento da teoria que dar suporte as análises estatísticas de um experimento em blocos completos casualizados com repetições dentro dos blocos. Neste sentido, faz-se necessário apresentar um desenho no espaço desse tipo de experimento, bem como, sugerir a construção de uma tabela para recolhimento das observações.

2.1 Um possível desenho do experimento no campo

Para ilustrar a localização espacial das unidades experinetais levou-se em conta um experimento com $I = 3$ tratamentos, $J = 4$ blocos e $R = 2$ repetições dos tratamentos dentro de cada bloco. Com estas características o experimento poderá ter o seguinte desenho no campo:



2.2 Organização dos dados experimentais

A organização dos dados coletados no campo em tabelas apropriadas, facilitará o tratamento estatístico posteriormente. Neste sentido, a tabela a seguir é uma sugestão para esta finalidade.

Tabela 1: Tabela para recolhimento dos dados no campo.

Tratamento	Repetição	Bloco				Soma
		1	2	...	J	
1	1	y_{111}	y_{121}	...	y_{1J1}	$y_{1.1}$
	2	y_{112}	y_{122}	...	y_{1J2}	$y_{1.2}$
	⋮	⋮	⋮		⋮	⋮
	R	y_{11R}	y_{21R}	...	y_{1JR}	$y_{1.R}$
			$y_{11.}$	$y_{12.}$...	$y_{1J.}$
2	1	y_{211}	y_{221}	...	y_{2J1}	$y_{2.1}$
	2	y_{212}	y_{222}	...	y_{2J2}	$y_{2.2}$
	⋮	⋮	⋮		⋮	⋮
	R	y_{21R}	y_{22R}	...	y_{2JR}	$y_{2.R}$
			$y_{21.}$	$y_{22.}$...	$y_{2J.}$
⋮	⋮	⋮	⋮	...	⋮	⋮
⋮	⋮	⋮	⋮	...	⋮	⋮
I	1	y_{I11}	y_{I21}	...	y_{IJ1}	$y_{I.1}$
	2	y_{I12}	y_{I22}	...	y_{IJ2}	$y_{I.2}$
	⋮	⋮	⋮		⋮	⋮
	R	y_{I1R}	y_{I2R}	...	y_{IJR}	$y_{I.R}$
			$y_{I1.}$	$y_{I2.}$...	$y_{IJ.}$
Soma		$y_{.1.}$	$y_{.2.}$...	$y_{.J.}$	$y_{...}$

Em que, y_{ijr} é a observação obtida da r -ésima unidade experimental que recebeu o tratamento I no bloco J , $y_{i..} = \sum_{j=1}^J \sum_{r=1}^R y_{ijr}$ é o total das JR observações que receberam o i -ésimo tratamento, $y_{.j.} = \sum_{i=1}^I \sum_{r=1}^R y_{ijr}$ é a soma das IR observações do j -ésimo bloco, $y_{ij.} = \sum_{r=1}^R y_{ijr}$ é a soma das R observações oriundas das unidades experimentais que receberam o tratamento I no bloco J e $y_{...} = \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R y_{ijr}$ é a soma de todas as observações.

2.3 O Modelo matemático

O modelo matemático adequado para descrever as observações de um experimento em blocos ao acaso com repetições dentro dos blocos é, conforme Barbin (1993), como segue:

$$y_{ijr} = \mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr}, \quad \begin{cases} i = 1, 2, \dots, I, \\ j = 1, 2, \dots, J, \\ r = 1, 2, \dots, R, \end{cases} \quad (2.1)$$

no qual:

y_{ijr} é a observação obtida da r -ésima unidade experimental do bloco j que recebeu o i -ésimo tratamento;

μ é a média geral;

τ_i é o efeito do tratamento i sobre a variável resposta, considerado fixo;

β_j é o efeito do bloco j sobre a variável resposta, também considerado fixo;

ϵ_{ij} e ε_{ijr} são respectivamente, erros atribuídos as unidades experimentais, entre e dentro dos blocos, ambos aleatórios, independentes e identicamente distribuídos como uma normal de médias zero e variâncias σ_ϵ^2 e σ^2 , respectivamente os quais serão denotados por:

$$\epsilon_{ij} \stackrel{iid}{\sim} N(0; \sigma_\epsilon^2) \quad \text{e} \quad \varepsilon_{ijr} \stackrel{iid}{\sim} N(0; \sigma^2).$$

Decorre das suposições acerca dos termos no modelo (2.1) que:

$$\begin{aligned} E(\mu) &= \mu; & E(\mu^2) &= \mu^2; & E(\tau_i) &= \tau_i; & E(\tau_i^2) &= \tau_i^2; & E(\beta_j) &= \beta_j; \\ E(\beta_j^2) &= \beta_j^2; & E(\epsilon_{ij}) &= 0; & E(\epsilon_{ij}^2) &= \sigma_\epsilon^2; & E(\varepsilon_{ijr}) &= 0; & E(\varepsilon_{ijr}^2) &= \sigma^2; \end{aligned}$$

$$\begin{aligned} E(\epsilon_{ij}\epsilon_{ks}) &= E(\epsilon_{ij})E(\epsilon_{ks}) = 0, \forall i \neq k \text{ ou } j \neq s; \\ E(\varepsilon_{ijr}\varepsilon_{ksv}) &= E(\varepsilon_{ijr})E(\varepsilon_{ksv}) = 0, \forall i \neq k, j \neq s \text{ ou } r \neq v; \\ E(\epsilon_{ij}\varepsilon_{ijr}) &= E(\epsilon_{ij})E(\varepsilon_{ijr}) = 0, \forall i, j, r. \end{aligned}$$

2.4 Estimação dos parâmetros, dos erros e das observações

O conjunto de dados observados num experimento em blocos ao acaso com repetições dentro dos blocos, podem ser escrito da seguinte forma:

$$\{y_{111}, \dots, y_{11R}, y_{121}, \dots, y_{12R}, \dots, \dots, y_{IJ1}, \dots, y_{IJR}\}$$

e podem ser representadas pelo modelo:

$$y_{ijr} = m + t_i + b_j + e_{ij} + u_{ijr} \quad \begin{cases} i = 1, 2, \dots, I \\ j = 1, 2, \dots, J \\ r = 1, 2, \dots, R \end{cases} \quad (2.2)$$

em que,

m é o estimador de μ , a média geral;

t_i é o estimador de τ_i , o efeito do tratamento i sobre a variável resposta;

b_j é o estimador de β_j , o efeito do bloco j sobre a variável resposta;

e_{ij} é o estimador de ϵ_{ij} , o erro entre blocos;

u_{ijr} é o estimador de ε_{ijr} , o erro dentro dos blocos.

Como foi dito no início, o método utilizado para encontrar os estimadores dos termos no modelo (2.2), foi o de *mínimos quadrados*, que consiste de encontrar os estimadores, de modo que torne mínima a soma dos quadrados dos erros dentro dos blocos.

Do modelo (2.2), tem-se que,

$$u_{ijr} = y_{ijr} - m - t_i - b_j - e_{ij} \implies u_{ijr}^2 = (y_{ijr} - m - t_i - b_j - e_{ij})^2$$

Somando-se para todas as observações, vem

$$Z = \sum_{ijr} u_{ijr}^2 = \sum_{ijr} (y_{ijr} - m - t_i - b_j - e_{ij})^2$$

Derivando-se parcialmente a função Z em relação a cada estimador, igualando-se a zero e explicitando cada um deles, vem:

$$\begin{aligned} \frac{\partial Z}{\partial m} &= 2 \sum_{ijr} (y_{ijr} - m - t_i - b_j)(-1) = 0 \\ &= y_{\dots} - IJRm - JR \sum_i t_i - IR \sum_j b_j = 0 \\ \therefore IJRm + JR \sum_i t_i + IR \sum_j b_j &= y_{\dots} \end{aligned} \quad (2.3)$$

$$\begin{aligned}
\frac{\partial Z}{\partial t_i} &= 2 \sum_{jr} (y_{ijr} - m - t_i - b_j - e_{ij})(-1) = 0 \\
&= y_{i..} - JRm - JRt_i - R \sum_j b_j - R \sum_j e_{ij} = 0 \\
\therefore JRm + JRt_i + R \sum_j b_j + R \sum_j e_{ij} &= y_{i..} \tag{2.4}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial Z}{\partial b_j} &= 2 \sum_{ir} (y_{ijr} - m - t_i - b_j - e_{ij})(-1) = 0 \\
&= y_{.j.} - IRm - R \sum_i t_i - IRb_j - R \sum_j e_{ij} = 0 \\
\therefore IRm + R \sum_i t_i + IRb_j + R \sum_j e_{ij} &= y_{.j.} \tag{2.5}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial Z}{\partial e_{ij}} &= 2 \sum_r (y_{ijr} - m - t_i - b_j - e_{ij})(-1) = 0 \\
&= y_{ij.} - Rm - Rt_i - Rb_j - Re_{ij} = 0 \\
\therefore Rm + Rt_i + Rb_j + Re_{ij} &= y_{ij.} \tag{2.6}
\end{aligned}$$

O sistema formado pelas Equações de (2.3) a (2.6) é conhecido na literatura por sistema de equações normais. Isto é,

$$\left\{ \begin{array}{l}
IJRm + JR \sum_i t_i + IR \sum_j b_j = y_{...} \\
JRm + JRt_i + R \sum_j b_j + R \sum_j e_{ij} = y_{i..} \\
IRm + R \sum_i t_i + IRb_j + R \sum_j e_{ij} = y_{.j.} \\
Rm + Rt_i + Rb_j + Re_{ij} = y_{ij.}
\end{array} \right. \tag{2.7}$$

O sistema de equações em (2.7) é inconsistente, e para resolve-lo é preciso impor as seguintes restrições:

$$\sum_i t_i = \sum_j b_j = \sum_i e_{ij} = \sum_j e_{ij} = \sum_{ij} e_{ij} = 0.$$

Assim sendo, o sistema (2.7), fica:

$$\begin{cases} IJRm & = y_{..} \\ JRm + JRt_i & = y_{i..} \\ IRm + IRb_j & = y_{.j.} \\ Rm + Rt_i + Rb_j + Re_{ij} & = y_{ij.} \end{cases} \quad (2.8)$$

Resolvendo o sistema (2.8) obtém-se os estimadores de mínimos quadrados:

$$m = \bar{y}_{..}; \quad t_i = \bar{y}_{i..} - \bar{y}_{..}; \quad b_j = \bar{y}_{.j.} - \bar{y}_{..}; \quad e_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{..} \quad (2.9)$$

Além disso, tem-se que

$$\begin{aligned} \hat{y}_{ijr} &= m + t_i + b_j + e_{ij} \\ &= \bar{y}_{..} + \bar{y}_{i..} - \bar{y}_{..} + \bar{y}_{.j.} - \bar{y}_{..} + \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{..} \\ &= \bar{y}_{ij.} \end{aligned}$$

daí, tem-se que $\bar{y}_{ij.}$ é um estimador de mínimos quadrados de y_{ijr} .

Fazendo-se as devidas substituições na Equação (2.2), obtém-se o estimador do erro dentro dos blocos, isto é,

$$u_{ijr} = y_{ijr} - \hat{y}_{ijr} = (y_{ijr} - \bar{y}_{..}) - (\bar{y}_{i..} - \bar{y}_{..}) - (\bar{y}_{.j.} - \bar{y}_{..}) - (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{..}). \quad (2.10)$$

Um resumo dos resultados obtidos nesta seção é apresentado na Tabela 2 a seguir.

Tabela 2: Estimadores das características envolvidas no modelo matemático.

Características	Estimador
μ	$\hat{\mu} = m = \bar{y}_{..}$
τ_i	$\hat{\tau}_i = t_i = \bar{y}_{i..} - \bar{y}_{..}$
β_j	$\hat{\beta}_j = b_j = \bar{y}_{.j.} - \bar{y}_{..}$
ϵ_{ij}	$\hat{\epsilon}_{ij} = e_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{..}$
ϵ_{ijr}	$\hat{\epsilon}_{ijr} = u_{ijr} = y_{ijr} - \bar{y}_{ij.}$
y_{ijr}	$\hat{y}_{ijr} = \bar{y}_{ij.}$

2.5 Decomposição da variabilidade total

Elevando-se ao quadrado os dois lados da Equação (2.10) e somando-se para todas as observações, vem:

$$\begin{aligned}
\sum_{ijr} u_{ijr}^2 &= \sum_{ijr} [(y_{ijr} - \bar{y}_{...}) - (\bar{y}_{i..} - \bar{y}_{...}) - (\bar{y}_{.j.} - \bar{y}_{...}) - (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})]^2 \\
&= \underbrace{\sum_{ijr} (y_{ijr} - \bar{y}_{...})^2}_{(1)} + \underbrace{JR \sum_i (\bar{y}_{i..} - \bar{y}_{...})^2}_{(2)} + \underbrace{IR \sum_j (\bar{y}_{.j.} - \bar{y}_{...})^2}_{(3)} \\
&\quad + \underbrace{R \sum_{ij} (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2}_{(4)} - \underbrace{2 \sum_{ijr} (y_{ijr} - \bar{y}_{...})(\bar{y}_{i..} - \bar{y}_{...})}_{(5)} \\
&\quad - \underbrace{2 \sum_{ijr} (y_{ijr} - \bar{y}_{...})(\bar{y}_{.j.} - \bar{y}_{...})}_{(6)} - \underbrace{2 \sum_{ijr} (y_{ijr} - \bar{y}_{...})(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})}_{(7)} \\
&\quad + \underbrace{2R \sum_{ij} (\bar{y}_{i..} - \bar{y}_{...})(\bar{y}_{.j.} - \bar{y}_{...})}_{(8)} + \underbrace{2R \sum_{ij} (\bar{y}_{i..} - \bar{y}_{...})(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})}_{(9)} \\
&\quad + \underbrace{2R \sum_{ij} (\bar{y}_{.j.} - \bar{y}_{...})(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})}_{(10)} \tag{2.11}
\end{aligned}$$

Desenvolvido-se algebricamente os termos de (1) a (10) da expressão (2.11), obtém-se os seguintes resultados:

$$\begin{aligned}
(1) &= \sum_{ijr} (y_{ijr} - \bar{y}_{...})^2 = \underbrace{\sum_{ijr} y_{ijr}^2}_{SQ_{Total}} - C, \quad \text{em que } C = \frac{y_{...}^2}{IJR}; \\
(2) &= JR \sum_i (\bar{y}_{i..} - \bar{y}_{...})^2 = \underbrace{\frac{1}{JR} \sum_i y_{i..}^2}_{SQ_{Tratamento}} - C; \\
(3) &= IR \sum_j (\bar{y}_{.j.} - \bar{y}_{...})^2 = \underbrace{\frac{1}{IR} \sum_j y_{.j.}^2}_{SQ_{Blocos}} - C; \\
(4) &= R \sum_{ij} (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\
&= \underbrace{\left(\frac{1}{R} \sum_{ij} y_{ij.}^2 - C \right)}_{SQ_{Parcelas}} - \underbrace{\left(\frac{1}{JR} \sum_i y_{i..}^2 - C \right)}_{SQ_{Tratamento}} - \underbrace{\left(\frac{1}{IR} \sum_j y_{.j.}^2 - C \right)}_{SQ_{Blocos}}; \\
&\quad \underbrace{\hspace{10em}}_{SQ_{Erro\ Entre}}
\end{aligned}$$

$$\begin{aligned}
(5) &= \sum_{ijr} (y_{ijr} - \bar{y}_{...})(\bar{y}_{i..} - \bar{y}_{...}) = \underbrace{\frac{1}{JR} \sum_i y_{i..}^2 - C}_{SQ_{Tratamento}}; \\
(6) &= \sum_{ijr} (y_{ijr} - \bar{y}_{...})(\bar{y}_{.j.} - \bar{y}_{...}) = \underbrace{\frac{1}{IR} \sum_j y_{.j.}^2 - C}_{SQ_{Blocos}}; \\
(7) &= \sum_{ijr} (y_{ijr} - \bar{y}_{...})(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}) \\
&= \underbrace{\left(\frac{1}{R} \sum_{ij} y_{ij.}^2 - C\right)}_{SQ_{Parcelas}} - \underbrace{\left(\frac{1}{JR} \sum_i y_{i..}^2 - C\right)}_{SQ_{Tratamento}} - \underbrace{\left(\frac{1}{IR} \sum_j y_{.j.}^2 - C\right)}_{SQ_{Blocos}}; \\
(8) &= R \sum_{ij} (\bar{y}_{i..} - \bar{y}_{...})(\bar{y}_{.j.} - \bar{y}_{...}) = R \sum_i (\bar{y}_{i..} - \bar{y}_{...}) \sum_j (\bar{y}_{.j.} - \bar{y}_{...}) \\
&= R \sum_i \left(\frac{y_{i..}}{JR} - \bar{y}_{...}\right) \sum_j \left(\frac{y_{.j.}}{IR} - \bar{y}_{...}\right) \\
&= \left(R \frac{y_{...}}{JR} - IR \frac{y_{...}}{IJR}\right) \left(R \frac{y_{...}}{IR} - RJ \frac{y_{...}}{IJR}\right) \\
&= \left(\frac{y_{...}}{J} - \frac{y_{...}}{J}\right) \left(\frac{y_{...}}{I} - \frac{y_{...}}{I}\right) = 0 \\
(9) &= R \sum_{ij} (\bar{y}_{i..} - \bar{y}_{...})(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}) \\
&= R \sum_i \left(\frac{y_{i..}}{JR} - \frac{y_{...}}{IJR}\right) \sum_j \left(\frac{y_{ij.}}{R} - \frac{y_{i..}}{JR} - \frac{y_{.j.}}{IR} + \frac{y_{...}}{IJR}\right) \\
&= \sum_i \left(\frac{y_{i..}}{J} - \frac{y_{...}}{IJ}\right) \left(\frac{y_{i..}}{R} - J \frac{y_{i..}}{JR} - \frac{y_{...}}{IR} + J \frac{y_{...}}{IJR}\right) = 0 \\
(10) &= R \sum_{ij} (\bar{y}_{.j.} - \bar{y}_{...})(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}) \\
&= R \sum_j \left(\frac{y_{.j.}}{IR} - \frac{y_{...}}{IJR}\right) \sum_i \left(\frac{y_{ij.}}{R} - \frac{y_{i..}}{JR} - \frac{y_{.j.}}{IR} + \frac{y_{...}}{IJR}\right) \\
&= \sum_j \left(\frac{y_{.j.}}{I} - \frac{y_{...}}{IJ}\right) \left(\frac{y_{.j.}}{R} - \frac{y_{...}}{JR} - I \frac{y_{.j.}}{IR} + I \frac{y_{...}}{IJR}\right) = 0
\end{aligned}$$

Substituindo-se estes resultados em (2.11), vem

$$\begin{aligned}
\sum_{ijr} u_{ijr}^2 &= \left(\sum_{ijr} y_{ijr}^2 - C\right) + \left(\frac{1}{JR} \sum_i y_{i..}^2 - C\right) + \left(\frac{1}{IR} \sum_j y_{.j.}^2 - C\right) \\
&\quad - \left[\left(\frac{1}{R} \sum_{ij} y_{ij.}^2 - C\right) - \left(\frac{1}{JR} \sum_i y_{i..}^2 - C\right) - \left(\frac{1}{IR} \sum_j y_{.j.}^2 - C\right)\right] \\
&\quad - 2 \left(\frac{1}{JR} \sum_i y_{i..}^2 - C\right) - 2 \left(\frac{1}{IR} \sum_j y_{.j.}^2 - C\right)
\end{aligned}$$

$$\begin{aligned}
& -2 \left[\left(\frac{1}{R} \sum_{ij} y_{ij}^2 - C \right) - \left(\frac{1}{JR} \sum_i y_{i..}^2 - C \right) - \left(\frac{1}{IR} \sum_j y_{.j}^2 - C \right) \right] \\
& + 2 \times 0 + 2 \times 0 + 2 \times 0 \\
\sum_{ijr} u_{ijr}^2 & = \left(\sum_{ijr} y_{ijr}^2 - C \right) - \left(\frac{1}{JR} \sum_i y_{i..}^2 - C \right) - \left(\frac{1}{IR} \sum_j y_{.j}^2 - C \right) \\
& - \left[\left(\frac{1}{R} \sum_{ij} y_{ij}^2 - C \right) - \left(\frac{1}{JR} \sum_i y_{i..}^2 - C \right) - \left(\frac{1}{IR} \sum_j y_{.j}^2 - C \right) \right]
\end{aligned}$$

ou,

$$\begin{aligned}
\underbrace{\left(\sum_{ijr} y_{ijr}^2 - C \right)}_{SQ_{Total}} & = \underbrace{\sum_{ijr} u_{ijr}^2}_{SQ_{Erro Dentro}} + \underbrace{\left(\frac{1}{JR} \sum_i y_{i..}^2 - C \right)}_{SQ_{Tratamento}} + \underbrace{\left(\frac{1}{IR} \sum_j y_{.j}^2 - C \right)}_{SQ_{Blocos}} \\
& + \underbrace{\left(\frac{1}{R} \sum_{ij} y_{ij}^2 - C \right)}_{SQ_{Parcelas}} - \underbrace{\left(\frac{1}{JR} \sum_i y_{i..}^2 - C \right)}_{SQ_{Tratamento}} - \underbrace{\left(\frac{1}{IR} \sum_j y_{.j}^2 - C \right)}_{SQ_{Blocos}} \\
& \underbrace{\hspace{10em}}_{SQ_{Erro Entre}}
\end{aligned}$$

Portanto, a soma dos quadrados total é decomposta em *quatro* partes, a saber,

$$SQ_{Total} = SQ_{Blocos} + SQ_{Tratamentos} + SQ_{Erro Entre} + SQ_{Erro Dentro}.$$

Na prática, a $SQ_{Erro Dentro}$ é calculada da seguinte maneira:

$$\underbrace{\sum_{ijr} u_{ijr}^2}_{SQ_{Erro Dentro}} = \underbrace{\left(\sum_{ijr} y_{ijr}^2 - C \right)}_{SQ_{Total}} - \underbrace{\left(\frac{1}{R} \sum_{ij} y_{ij}^2 - C \right)}_{SQ_{Parcelas}}. \quad (2.12)$$

Os resultados da decomposição da variabilidade total é organizada na Tabela 3, a qual é conhecida na literatura por *Tabela da Análise da Variância - ANOVA*.

Tabela 3: Tabela da Análise da Variância - ANOVA

F.V.	G.L.	S.Q.	Q.M.	F
Tratamento	$I - 1$	$\frac{1}{JR} \sum_i y_{i..}^2 - C$	QM_{Trat}	
Blocos	$J - 1$	$\frac{1}{IR} \sum_j y_{.j.}^2 - C$	QM_{Blocos}	
Erro Entre Parcelas	$(I - 1)(J - 1)$	$SQ_{Parc} - SQ_{Trat} - SQ_{Blocos}$	$QM_{Erro Entre}$	
Parcelas	$(IJ - 1)$	$\frac{1}{R} \sum_{ij} y_{ij.}^2 - C$	-	
Erro Dentro Parcelas	$IJ(R - 1)$	$SQ_{Total} - SQ_{Parc}$	$QM_{Erro Entre}$	
Total	$IJR - 1$	$\sum_{ijr} y_{ijr}^2 - C$	-	

em que, os resultados da coluna 4, da Tabela 3, referentes aos Quadrados Médios (Q.M.), são obtidos por meio da divisão dos elementos da coluna 3, (S.Q.) pelos respectivos elementos da coluna 2, (G.L.). Os elementos da coluna 5, (F), serão discutidos e apresentados posteriormente.

2.6 Distribuição de probabilidade dos estimadores

Nesta seção será estudada as distribuições de probabilidade dos estimadores, bem como algumas propriedades destes. Os elementos a seguir ajudarão nas demonstrações das características associadas às distribuições de probabilidade dos estimadores a serem desenvolvidas.

$$y_{ijr} = \mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr}$$

e que,

$$\begin{aligned} E(\epsilon_{ij}) &= 0, & E(\epsilon_{ij}^2) &= \sigma_\epsilon^2 \implies \epsilon_{ij} \stackrel{iid}{\sim} N(0; \sigma_\epsilon^2), \\ E(\varepsilon_{ijr}) &= 0, & E(\varepsilon_{ijr}^2) &= \sigma^2 \implies \varepsilon_{ijr} \stackrel{iid}{\sim} N(0; \sigma^2), \end{aligned}$$

consequentemente, tem-se:

$$E(y_{ijr}) = \mu + \tau_i + \beta_j, \quad Var(y_{ijr}) = \sigma_\epsilon^2 + \sigma^2$$

e segue que

$$y_{ijr} \sim N(\mu + \tau_i + \beta_j; \sigma_\epsilon^2 + \sigma^2).$$

2.6.1 Distribuição de probabilidade de $\bar{y}_{...}$, o estimador de μ .

$\bar{y}_{...} = \frac{1}{IJR} \sum_{ijr} y_{ijr} = \frac{1}{IJR} (y_{111} + y_{112} + \dots + y_{IJR})$, que é uma combinação linear dos y_{ijr} , os quais seguem distribuição normal. Como sabe-se que combinação linear de variáveis normais é também normal, então $\bar{y}_{...}$ segue uma distribuição normal;

Mas, sendo

$$\begin{aligned} \bar{y}_{...} &= \frac{1}{IJR} \sum_{ijr} y_{ijr} = \frac{1}{IJR} \sum_{ijr} (\mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr}) \\ &= \frac{1}{IJR} \left(IJR\mu + JR \sum_i \tau_i + IR \sum_j \beta_j + R \sum_{ij} \epsilon_{ij} + \sum_{ijr} \varepsilon_{ijr} \right) \\ &= \mu + \frac{1}{IJ} \sum_{ij} \epsilon_{ij} + \frac{1}{IJR} \sum_{ijr} \varepsilon_{ijr}, \end{aligned}$$

as características da distribuição de $\bar{y}_{...}$ serão determinadas como segue:

$$\begin{aligned} E(\bar{y}_{...}) &= \frac{1}{IJR} E \left(\sum_{ijr} y_{ijr} \right) = \frac{1}{IJR} \sum_{ijr} E(y_{ijr}) = \frac{1}{IJR} \sum_{ijr} (\mu + \tau_i + \beta_j) \\ &= \frac{1}{IJR} \left(IJR\mu + JR \sum_i \tau_i + IR \sum_j \beta_j \right), \end{aligned}$$

mas por definição μ é a média geral e, portanto, cumpre-se que

$$\sum_i \tau_i = \sum_j \beta_j = 0,$$

logo,

$$E(\bar{y}_{...}) = \mu. \quad (2.13)$$

Além disso,

$$\begin{aligned} Var(\bar{y}_{...}) &= E[\bar{y}_{...} - E(\bar{y}_{...})]^2 = E[\bar{y}_{...} - \mu]^2 \\ &= E \left[\mu + \frac{1}{IJ} \sum_{ij} \epsilon_{ij} + \frac{1}{IJR} \sum_{ijr} \varepsilon_{ijr} - \mu \right]^2 \\ &= E \left[\frac{1}{IJ} \sum_{ij} \epsilon_{ij} + \frac{1}{IJR} \sum_{ijr} \varepsilon_{ijr} \right]^2 \\ &= E \left[\frac{1}{IJ} (\epsilon_{11} + \dots + \epsilon_{IJ}) + \frac{1}{IJR} (\varepsilon_{111} + \dots + \varepsilon_{IJR}) \right]^2 \end{aligned}$$

$$\begin{aligned}
&= E \left[\frac{1}{I^2 J^2} (\epsilon_{11}^2 + \dots + \epsilon_{IJ}^2 + dp) \right. \\
&\quad \left. + \frac{1}{I^2 J^2 R^2} (\epsilon_{111}^2 + \dots + \epsilon_{IJR}^2 + dp) + dp \right] \\
&= \frac{1}{I^2 J^2} (\sigma_\epsilon^2 + \dots + \sigma_\epsilon^2 + 0) + \frac{1}{I^2 J^2 R^2} (\sigma^2 + \dots + \sigma^2 + 0) + 0 \\
&= \frac{1}{I^2 J^2} IJ \sigma_\epsilon^2 + \frac{1}{I^2 J^2 R^2} IJR \sigma^2 = \frac{\sigma_\epsilon^2}{IJ} + \frac{\sigma^2}{IJR}
\end{aligned}$$

e segue que

$$Var(\bar{y}_{...}) = \frac{1}{IJR} (\sigma^2 + R\sigma_\epsilon^2). \quad (2.14)$$

Portanto,

$$\bar{y}_{...} \sim N \left(\mu; \frac{1}{IJR} (\sigma^2 + R\sigma_\epsilon^2) \right). \quad (2.15)$$

Obs.: dp = duplos produtos da equação.

2.6.2 Distribuição de probabilidade da correção para média, C .

A distribuição de probabilidade de C é obtida do seguinte modo:

Sendo,

$$\bar{y}_{...} \sim N \left(\mu; \frac{1}{IJR} (\sigma^2 + R\sigma_\epsilon^2) \right) \implies \frac{\bar{y}_{...} - \mu}{\sqrt{\frac{\sigma^2 + R\sigma_\epsilon^2}{IJR}}} \sim N(0; 1).$$

Assim, sob $H_0 : \mu = 0$

$$\frac{\bar{y}_{...} - 0}{\sqrt{\frac{\sigma^2 + R\sigma_\epsilon^2}{IJR}}} \sim N(0; 1) \implies \frac{\bar{y}_{...}^2}{\frac{\sigma^2 + R\sigma_\epsilon^2}{IJR}} = \frac{IJR \frac{y_{...}}{IJR} \frac{y_{...}}{IJR}}{\sigma^2 + R\sigma_\epsilon^2} = \frac{\frac{y_{...}^2}{IJR}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{(1)}^2.$$

Isto é, a estatística $\frac{C}{\sigma^2 + R\sigma_\epsilon^2}$ segue uma distribuição de qui-quadrado com 1 grau de liberdade e será denotado por:

$$\frac{C}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{(1)}^2, \quad (2.16)$$

em que $C = \frac{y_{...}^2}{IJR}$.

2.6.3 Distribuição de probabilidade de t_i , o estimador de τ_i .

Para se obter a distribuição de probabilidade do estimador do efeito do i -ésimo tra-

tamento pode ser usado o seguinte procedimento:

$t_i = \bar{y}_{i..} - \bar{y}_{...} = \frac{1}{JR} \sum_{jr} y_{ijr} - \frac{1}{IJR} \sum_{ijr} y_{ijr}$, que é uma combinação linear dos y_{ijr} que segue uma distribuição normal e, portanto, t_i é normal.

Além disso,

$$\begin{aligned}
 t_i &= \frac{1}{JR} \left(JR\mu + JR\tau_i + R \sum_j \beta_j + R \sum_j \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr} \right) \\
 &\quad - \frac{1}{IJR} \left(IJR\mu + JR \sum_i \tau_i + IR \sum_j \beta_j + R \sum_{ij} \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr} \right) \\
 &= \left(\mu + \tau_i + \frac{1}{J} \sum_j \epsilon_{ij} + \frac{1}{JR} \sum_{jr} \epsilon_{ijr} \right) - \left(\mu + \frac{1}{IJ} \sum_{ij} \epsilon_{ij} + \frac{1}{IJR} \sum_{ijr} \epsilon_{ijr} \right) \\
 &= \tau_i + \frac{1}{J} \sum_j \epsilon_{ij} + \frac{1}{JR} \sum_{jr} \epsilon_{ijr} - \frac{1}{IJ} \sum_{ij} \epsilon_{ij} - \frac{1}{IJR} \sum_{ijr} \epsilon_{ijr}.
 \end{aligned}$$

Portanto,

$$E(t_i) = \tau_i. \quad (2.17)$$

Além disso, tem-se que

$$\begin{aligned}
 Var(t_i) &= E[t_i - E(t_i)]^2 \\
 &= E\left[\tau_i + \frac{1}{J} \sum_j \epsilon_{ij} + \frac{1}{JR} \sum_{jr} \epsilon_{ijr} - \frac{1}{IJ} \sum_{ij} \epsilon_{ij} - \frac{1}{IJR} \sum_{ijr} \epsilon_{ijr} - \tau_i\right]^2 \\
 &= E\left[\frac{1}{J} \sum_j \epsilon_{ij} + \frac{1}{JR} \sum_{jr} \epsilon_{ijr} - \frac{1}{IJ} \sum_{ij} \epsilon_{ij} - \frac{1}{IJR} \sum_{ijr} \epsilon_{ijr}\right]^2 \\
 &= E\left[\frac{1}{J^2} \left(\sum_j \epsilon_{ij}\right)^2 + \frac{1}{J^2 R^2} \left(\sum_{jr} \epsilon_{ijr}\right)^2 + \frac{1}{I^2 J^2} \left(\sum_{ij} \epsilon_{ij}\right)^2\right. \\
 &\quad \left. + \frac{1}{I^2 J^2 R^2} \left(\sum_{ijr} \epsilon_{ijr}\right)^2 + \frac{2}{J^2 R} \left(\sum_j \epsilon_{ij}\right) \left(\sum_{jr} \epsilon_{ijr}\right) - \frac{2}{IJ^2} \left(\sum_j \epsilon_{ij}\right) \left(\sum_{ij} \epsilon_{ij}\right)\right. \\
 &\quad \left. - \frac{2}{IJ^2 R} \left(\sum_j \epsilon_{ij}\right) \left(\sum_{ijr} \epsilon_{ijr}\right) - \frac{2}{IJ^2 R} \left(\sum_{jr} \epsilon_{ijr}\right) \left(\sum_{ij} \epsilon_{ij}\right)\right. \\
 &\quad \left. - \frac{2}{IJ^2 R^2} \left(\sum_{ijr} \epsilon_{ijr}\right) \left(\sum_{ijr} \epsilon_{ijr}\right) + \frac{2}{I^2 J^2 R} \left(\sum_{ij} \epsilon_{ij}\right) \left(\sum_{ijr} \epsilon_{ijr}\right)\right] \\
 &= \frac{\sigma_\epsilon^2}{J} + \frac{\sigma^2}{JR} + \frac{\sigma_\epsilon^2}{IJ} + \frac{\sigma^2}{IJR} - 2\frac{\sigma_\epsilon^2}{IJ} - 2\frac{\sigma^2}{IJR} = \frac{\sigma_\epsilon^2}{J} + \frac{\sigma^2}{JR} - \frac{\sigma_\epsilon^2}{IJ} - \frac{\sigma^2}{IJR} \\
 &= \frac{1}{IJR} (IR\sigma_\epsilon^2 - R\sigma_\epsilon^2 + I\sigma^2 - \sigma^2) = \frac{1}{IJR} [R(I-1)\sigma_\epsilon^2 + (I-1)\sigma^2] \\
 &= \frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2).
 \end{aligned}$$

Portanto,

$$\text{Var}(t_i) = \frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2). \quad (2.18)$$

Assim sendo, conclui-se que

$$t_i \sim N\left(\tau_i; \frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2)\right). \quad (2.19)$$

Ou seja, t_i , tem distribuição normal com média τ_i e variância $\frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2)$, para todo $i = 1, 2, \dots, I$.

2.6.4 Distribuição de probabilidade da SQ_{Trat} .

A partir dos resultados da subseção anterior resumida na expressão (2.19) deduz-se que,

$$Z = \frac{t_i - \tau_i}{\sqrt{\frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2)}} = \frac{\bar{y}_{i..} - \bar{y}... - \tau_i}{\sqrt{\frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2)}} \sim N(0; 1)$$

e sob a hipótese de que os tratamentos não têm efeitos sobre a variável resposta, Y , isto é, sob $H_0 : \tau_1 = \tau_2 = \dots = \tau_I = 0$, então,

$$\frac{\bar{y}_{i..} - \bar{y}...}{\sqrt{\frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2)}} \sim N(0; 1) \implies \frac{(\bar{y}_{i..} - \bar{y}...)^2}{\frac{(I-1)}{IJR} (\sigma^2 + R\sigma_\epsilon^2)} \sim \chi_{(1)}^2$$

e

$$\begin{aligned} \frac{JR(\bar{y}_{i..} - \bar{y}...)^2}{\frac{(I-1)}{I} (\sigma^2 + R\sigma_\epsilon^2)} &\sim \chi_{(1)}^2 \implies \frac{JR(\bar{y}_{i..} - \bar{y}...)^2}{\sigma^2 + R\sigma_\epsilon^2} \sim \frac{(I-1)}{I} \chi_{(1)}^2 \implies \\ \frac{JR \sum_i (\bar{y}_{i..} - \bar{y}...)^2}{\sigma^2 + R\sigma_\epsilon^2} &\sim (I-1) \chi_{(1)}^2. \end{aligned}$$

Portanto,

$$\frac{JR \sum_i (\bar{y}_{i..} - \bar{y}...)^2}{\sigma^2 + R\sigma_\epsilon^2} = \frac{SQ_{Trat}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{(I-1)}^2. \quad (2.20)$$

Em palavras, a soma de quadrados devida aos tratamentos dividida por $\sigma^2 + R\sigma_\epsilon^2$ é distribuída como uma qui-quadrado com $(I-1)$ graus de liberdade.

2.6.5 Distribuição de probabilidade de b_j , o estimador de β_j .

Para se obter a distribuição de probabilidade do estimador do efeito do j -ésimo bloco pode ser usado o seguinte procedimento:

$b_j = \bar{y}_{.j} - \bar{y}_{...} = \frac{1}{IR} \sum_{ir} y_{ijr} - \frac{1}{IJR} \sum_{ijr} y_{ijr}$, é uma combinação linear dos y_{ijr} os quais seguem uma distribuição normal portanto, b_j também é normal;

Além disso,

$$\begin{aligned}
 b_j &= \frac{1}{IR} \left(IR\mu + R \sum_{i \neq j} \tau_i + IR\beta_j + R \sum_i \epsilon_{ij} + \sum_{ijr} \varepsilon_{ijr} \right) \\
 &\quad - \frac{1}{IJR} \left(IJR\mu + JR \sum_{i \neq j} \tau_i + IR \sum_{j \neq j} \beta_j + R \sum_{ij} \epsilon_{ij} + \sum_{ijr} \varepsilon_{ijr} \right) \\
 &= \mu + \beta_j + \frac{1}{I} \sum_i \epsilon_{ij} + \frac{1}{IR} \sum_{ir} \varepsilon_{ijr} - \mu - \frac{1}{IJ} \sum_{ij} \epsilon_{ij} - \frac{1}{IJR} \sum_{ijr} \varepsilon_{ijr} \\
 &= \beta_j + \frac{1}{I} \sum_i \epsilon_{ij} + \frac{1}{IR} \sum_{ir} \varepsilon_{ijr} - \frac{1}{IJ} \sum_{ij} \epsilon_{ij} - \frac{1}{IJR} \sum_{ijr} \varepsilon_{ijr}
 \end{aligned}$$

Portanto

$$E(b_j) = \beta_j. \quad (2.21)$$

Além disso, tem-se que

$$\begin{aligned}
 Var(b_j) &= E[b_j - E(b_j)]^2 \\
 &= E \left[\beta_j + \frac{1}{I} \sum_i \epsilon_{ij} + \frac{1}{IR} \sum_{ir} \varepsilon_{ijr} - \frac{1}{IJ} \sum_{ij} \epsilon_{ij} - \frac{1}{IJR} \sum_{ijr} \varepsilon_{ijr} - \beta_j \right]^2 \\
 &= E \left[\frac{1}{I^2} \left(\sum_i \epsilon_{ij} \right)^2 + \frac{1}{I^2 R^2} \left(\sum_{ir} \varepsilon_{ijr} \right)^2 + \frac{1}{I^2 J^2} \left(\sum_{ij} \epsilon_{ij} \right)^2 \right. \\
 &\quad \left. + \frac{1}{I^2 J^2 R^2} \left(\sum_{ijr} \varepsilon_{ijr} \right)^2 + dp - \frac{2}{I^2 J} \left(\sum_i \epsilon_{ij} \right) \left(\sum_{ij} \epsilon_{ij} \right) - dp \right. \\
 &\quad \left. - dp + \frac{2}{I^2 J R^2} \left(\sum_{ir} \varepsilon_{ijr} \right) \left(\sum_{ijr} \varepsilon_{ijr} \right) + dp \right] \\
 &= \frac{\sigma_\epsilon^2}{I} + \frac{\sigma^2}{IR} + \frac{\sigma_\epsilon^2}{IJ} + \frac{\sigma^2}{IJR} - 2 \frac{\sigma_\epsilon^2}{IJ} - 2 \frac{\sigma^2}{IJR} \\
 &= \frac{\sigma_\epsilon^2}{I} + \frac{\sigma^2}{IR} - \frac{\sigma_\epsilon^2}{IJ} - \frac{\sigma^2}{IJR} \\
 &= \frac{1}{IJR} (JR\sigma_\epsilon^2 + J\sigma^2 - R\sigma_\epsilon^2 - \sigma^2) = \frac{1}{IJR} (R(J-1)\sigma_\epsilon^2 + (J-1)\sigma^2)
 \end{aligned}$$

Portanto

$$\text{Var}(b_j) = \frac{J-1}{IJR}(\sigma^2 + R\sigma_\epsilon^2). \quad (2.22)$$

Com isso, obtém-se

$$b_j \sim N\left(\beta_j; \frac{J-1}{IJR}(\sigma^2 + R\sigma_\epsilon^2)\right). \quad (2.23)$$

Ou seja, o estimador b_j do efeito do bloco j segue uma distribuição normal com média β_j e variância $\frac{J-1}{IJR}(\sigma^2 + R\sigma_\epsilon^2)$.

2.6.6 Distribuição de probabilidade da SQ_{Blocos}

Como já se sabe, $b_j \sim N\left(\beta_j; \frac{J-1}{IJR}(\sigma^2 + R\sigma_\epsilon^2)\right)$. Portanto, é possível deduzir que

$$\frac{b_j - E(b_j)}{\sqrt{\text{Var}(b_j)}} = \frac{\bar{y}_{.j} - \bar{y}_{...} - \beta_j}{\sqrt{\frac{J-1}{IJR}(\sigma^2 + R\sigma_\epsilon^2)}} \sim N(0; 1).$$

Assim, sob $H_0 : \beta_1 = \beta_2 = \dots = \beta_j = 0$,

$$\begin{aligned} \frac{\bar{y}_{.j} - \bar{y}_{...}}{\sqrt{\frac{J-1}{IJR}(\sigma^2 + R\sigma_\epsilon^2)}} &\sim N(0; 1) \implies \frac{(\bar{y}_{.j} - \bar{y}_{...})^2}{\frac{J-1}{IJR}(\sigma^2 + R\sigma_\epsilon^2)} \sim \chi_{(1)}^2 \implies \\ \frac{IR(\bar{y}_{.j} - \bar{y}_{...})^2}{\sigma^2 + R\sigma_\epsilon^2} &\sim \frac{(J-1)}{J} \chi_{(1)}^2 \end{aligned}$$

Portanto,

$$\frac{IR \sum_j (\bar{y}_{.j} - \bar{y}_{...})^2}{\sigma^2 + R\sigma_\epsilon^2} = \frac{SQ_{\text{Blocos}}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{(J-1)}^2 \quad (2.24)$$

ou seja, a soma de quadrados de blocos dividida por $(\sigma^2 + R\sigma_\epsilon^2)$ segue uma distribuição de qui-quadrado com $(J-1)$ graus de liberdade.

2.6.7 Distribuições de probabilidade de m_i , o estimador de $\mu_i = \mu + \tau_i$, (a média do tratamento i) e de m_j , o estimador de $\mu_j = \mu + \beta_j$, (a média do bloco j).

Sabe-se que m_i , o estimador da média do i -ésimo tratamento, $\mu_i = \mu + \tau_i$, é definido por

$$m_i = m + t_i = \bar{y}_{...} + \bar{y}_{i..} - \bar{y}_{...} = \bar{y}_{i..} = \frac{1}{JR} \sum_{j,r} y_{ijr}$$

que é uma combinação linear dos y_{ijk} 's os quais seguem uma distribuição normal e, portanto, m_i também é normal. Continuando-se o desenvolvimento algébrico de m_i , vem

$$\begin{aligned} m_i &= \frac{1}{JR} \sum_{j,r} y_{ijr} = \frac{1}{JR} \left(JR\mu + JR\tau_i + R \sum_j \beta_j + R \sum_j \epsilon_{ij} + \sum_{j,r} \epsilon_{ijr} \right) \\ &= \mu_i + \frac{1}{J} \sum_j \epsilon_{ij} + \frac{1}{JR} \sum_{j,r} \epsilon_{ijr}. \end{aligned} \quad (2.25)$$

O valor esperado de m_i é então

$$E(m_i) = E \left(\mu_i + \frac{1}{J} \sum_j \epsilon_{ij} + \frac{1}{JR} \sum_{j,r} \epsilon_{ijr} \right) = \mu_i \quad (2.26)$$

Por outro lado, a variância de m_i pode ser obtida como,

$$\begin{aligned} Var(m_i) &= E[m_i - E(m_i)]^2 = E \left[\mu_i + \frac{1}{J} \sum_j \epsilon_{ij} + \frac{1}{JR} \sum_{j,r} \epsilon_{ijr} - \mu_i \right]^2 \\ &= E \left[\frac{1}{J^2} (\epsilon_{i1}^2 + \dots + \epsilon_{iJ}^2 + dp) + \frac{1}{J^2 R^2} (\epsilon_{i11}^2 + \dots + \epsilon_{iJR}^2 + dp) + odp \right] \\ &= \frac{J}{J^2} \sigma_\epsilon^2 + \frac{JR}{J^2 R^2} \sigma^2 = \frac{1}{JR} (\sigma^2 + R\sigma_\epsilon^2). \end{aligned} \quad (2.27)$$

Assim sendo, conclui-se que

$$m_i \stackrel{iid}{\sim} N \left(\mu_i, \frac{1}{JR} (\sigma^2 + R\sigma_\epsilon^2) \right), \quad i = 1, 2, \dots, I. \quad (2.28)$$

Daí, conclui-se que o estimador da média do i -ésimo tratamento segue uma distribuição normal de média μ_i e variância $\frac{1}{JR} (\sigma^2 + R\sigma_\epsilon^2)$.

Por procedimento análogo, obtém-se a distribuição de probabilidade de m_j , ou seja,

$$m_j \stackrel{iid}{\sim} N \left(\mu_j, \frac{1}{JR} (\sigma^2 + R\sigma_\epsilon^2) \right), \quad j = 1, 2, \dots, J. \quad (2.29)$$

E, conclui-se que o estimador da média do j -ésimo bloco segue uma distribuição normal com média μ_j e variância $\frac{1}{JR} (\sigma^2 + R\sigma_\epsilon^2)$.

2.6.8 Distribuições de probabilidade de contrastes de interesse

Dois resultados de grande interesse neste trabalho diz respeito as distribuições de probabilidade dos estimadores de combinações lineares das médias dos tratamentos ou dos

blocos, $\hat{\Psi}_h = \sum_{i=1}^I c_{hi} m_i$ ou $\hat{\Psi}_h = \sum_{j=1}^J c_{hj} m_j$, respectivamente, em que $\sum_i c_{hi} = \sum_j c_{hj} = 0$. Aqui, serão apresentadas as demonstrações relativas as combinações lineares de médias dos tratamentos e intuitivamente serão apresentados os resultados para uma combinação linear de médias dos blocos. Considere o estimador de uma combinação linear das médias dos tratamentos

$$\hat{\Psi}_h = \sum_{i=1}^I c_{hi} m_i = \sum_{i=1}^I c_{hi} \bar{y}_{i..} = \sum_i c_{hi} \frac{1}{JR} \sum_{j,r} y_{ijr}$$

que é uma combinação linear de variáveis normais e, portanto, $\hat{\Psi}_h$ também segue uma distribuição normal.

Continuando-se o desenvolvimento algébrico, vem

$$\begin{aligned} \hat{\Psi}_h &= \sum_{i=1}^I c_{hi} \frac{1}{JR} \sum_{j,r} y_{ijr} \\ &= \sum_{i=1}^I c_{hi} \frac{1}{JR} \left(JR\mu + JR\tau_i + R \sum_{j=1}^J \beta_j + R \sum_{j=1}^J \epsilon_{ij} + \sum_{j,r} \epsilon_{ijr} \right) \\ &= \sum_{i=1}^I c_{hi} \left(\mu_i + \frac{1}{J} \sum_{j=1}^J \epsilon_{ij} + \frac{1}{JR} \sum_{j,r} \epsilon_{ijr} \right). \end{aligned} \quad (2.30)$$

Daí, o valor esperado do estimador $\hat{\Psi}_h$, fica

$$E(\hat{\Psi}_h) = E \left[\sum_{i=1}^I c_{hi} \left(\mu_i + \frac{1}{J} \sum_{j=1}^J \epsilon_{ij} + \frac{1}{JR} \sum_{j,r} \epsilon_{ijr} \right) \right] = \sum_{i=1}^I c_{hi} \mu_i$$

Para obter a variância do estimador $\hat{\Psi}_h$ procedeu-se do seguinte modo

$$\begin{aligned} Var(\hat{\Psi}_h) &= E[\hat{\Psi}_h - E(\hat{\Psi}_h)]^2 \\ &= E \left[\sum_{i=1}^I c_{hi} \mu_i + \sum_{i=1}^I c_{hi} \frac{1}{J} \sum_{j=1}^J \epsilon_{ij} + \sum_{i=1}^I c_{hi} \frac{1}{JR} \sum_{j,r} \epsilon_{ijr} - \sum_{i=1}^I c_{hi} \mu_i \right]^2 \\ &= E \left[\sum_{i=1}^I c_{hi} \frac{1}{J} \sum_{j=1}^J \epsilon_{ij} + \sum_{i=1}^I c_{hi} \frac{1}{JR} \sum_{j,r} \epsilon_{ijr} \right]^2 \\ &= E \left[\frac{1}{J} \left(\sum_{i=1}^I c_{hi} \epsilon_{i1} + \cdots + \sum_{i=1}^I c_{hi} \epsilon_{iJ} \right) + \frac{1}{JR} \left(\sum_{i=1}^I c_{hi} \epsilon_{i11} + \cdots + \sum_{i=1}^I c_{hi} \epsilon_{iJR} \right) \right]^2 \\ &= \left[\frac{1}{J^2} \left(J\sigma_\epsilon^2 \sum_{i=1}^I c_{hi}^2 \right) + \frac{1}{J^2 R^2} \left(JR\sigma^2 \sum_{i=1}^I c_{hi}^2 \right) \right] \end{aligned}$$

$$= \left[\frac{1}{J}\sigma_\epsilon^2 + \frac{1}{JR}\sigma^2 \right] \sum_{i=1}^I c_{hi}^2 = \frac{1}{JR}(\sigma^2 + R\sigma_\epsilon^2) \sum_{i=1}^I c_{hi}^2. \quad (2.31)$$

Assim sendo, conclui-se que o estimador de um contraste das médias dos tratamentos segue uma distribuição normal com média $\Psi_h = \sum_{i=1}^I c_{hi}\mu_i$ e variância $\frac{K_h}{JR}(\sigma^2 + R\sigma_\epsilon^2)$, em que K_h é a soma dos quadrados dos coeficientes do contraste h , isto é, $K_h = \sum_{i=1}^I c_{hi}^2$ e será denotado por

$$\hat{\Psi}_h \sim N \left(\sum_{i=1}^I c_{hi}\mu_i, \frac{K_h}{JR}(\sigma^2 + R\sigma_\epsilon^2) \right) \quad (2.32)$$

Usando um procedimento análogo, demonstra-se que o estimador de um contraste das médias dos blocos segue uma distribuição normal com média $\Psi_h = \sum_{j=1}^J c_{hj}\mu_j$ e variância $\frac{K_h}{IR}(\sigma^2 + R\sigma_\epsilon^2)$ o qual será denotado como

$$\hat{\Psi}_h \sim N \left(\sum_{j=1}^J c_{hj}\mu_j, \frac{K_h}{IR}(\sigma^2 + R\sigma_\epsilon^2) \right), \quad (2.33)$$

em que K_h é a soma dos quadrados dos coeficientes do contraste h , isto é, $K_h = \sum_{j=1}^J c_{hj}^2$.

2.6.9 Distribuição de probabilidade da soma de quadrados da h -ésima combinação linear das médias dos tratamentos

Foi demonstrado que a distribuição de probabilidade de $\hat{\Psi}_h = \sum_{i=1}^I c_{hi}\bar{y}_{i..}$, o estimador do h -ésimo contraste de médias dos tratamentos, $\Psi_h = \sum_{i=1}^I c_{hi}\mu_i$ é, de acordo com a expressão (2.32), distribuída como

$$\hat{\Psi}_h \sim N \left(\sum_{i=1}^I c_{hi}\mu_i, \frac{K_h}{JR}(\sigma^2 + R\sigma_\epsilon^2) \right).$$

Usando resultados conhecidos da teoria de probabilidade, deduz-se que

$$\frac{\hat{\Psi}_h - E(\hat{\Psi}_h)}{\sqrt{\frac{K_h}{JR}(\sigma^2 + R\sigma_\epsilon^2)}} = \frac{\sum_{i=1}^I c_{hi}\bar{y}_{i..} - \sum_{i=1}^I c_{hi}\mu_i}{\sqrt{\frac{K_h}{JR}(\sigma^2 + R\sigma_\epsilon^2)}} \sim N(0, 1).$$

e, sob a hipóteses de que o contraste h é nulo, isto é, sob $H_0^{(\Psi)} : \Psi_h = \sum_{i=1}^I c_{hi}\mu_i = 0$, a

estatística

$$\frac{\sum_{i=1}^I c_{hi} \bar{y}_{i..}}{\sqrt{\frac{K_h}{JR}(\sigma^2 + R\sigma_\epsilon^2)}} \sim N(0, 1)$$

e segue que

$$\frac{\left(\sum_{i=1}^I c_{hi} \bar{y}_{i..}\right)^2}{\frac{K_h}{JR}(\sigma^2 + R\sigma_\epsilon^2)} = \frac{\left(\sum_{i=1}^I c_{hi} y_{i..}\right)^2}{JRK_h(\sigma^2 + R\sigma_\epsilon^2)} = \frac{SQ_{Contraste}}{\sigma^2 + R\sigma_\epsilon^2} = \frac{SQ(\hat{\Psi}_h)}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi^2_{(1)} \quad (2.34)$$

em que, $SQ_{Contraste} = SQ(\hat{\Psi}_h) = \frac{\left(\sum_{i=1}^I c_{hi} y_{i..}\right)^2}{JRK_h}$ e $K_h = \sum_{i=1}^I c_{hi}^2$.

2.6.10 Distribuição de probabilidade de $\hat{\epsilon}_{ij}$, o estimador de ϵ_{ij} .

Um outro resultado útil diz respeito a distribuição de probabilidade de $\hat{\epsilon}_{ij}$, o estimador do erro entre parcelas, o qual pode ser obtido utilizando-se o seguinte procedimento:

$$\hat{\epsilon}_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...} = \frac{1}{R} \sum_r y_{ijr} - \frac{1}{JR} \sum_{jr} y_{ijr} - \frac{1}{IR} \sum_i y_{ijr} + \frac{1}{IJR} \sum_{ijr} y_{ijr},$$

como pode-se observar, $\hat{\epsilon}_{ij}$ é uma combinação linear dos y_{ijr} os quais seguem distribuição normal. Portanto, $\hat{\epsilon}_{ij}$ também segue uma distribuição normal. É necessário agora, saber quais as características da distribuição de $\hat{\epsilon}_{ij}$.

$$\begin{aligned} E(\hat{\epsilon}_{ij}) &= \frac{1}{R} \sum_r E(\mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr}) - \frac{1}{JR} \sum_{jr} E(\mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr}) \\ &\quad - \frac{1}{IR} \sum_{ir} E(\mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr}) + \frac{1}{IJR} \sum_{ijr} E(\mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr}) \\ &= \frac{1}{R}(R\mu + R\tau_i + R\beta_j) - \frac{1}{JR}(JR\mu + JR\tau_i + R\beta_j) \\ &\quad - \frac{1}{IR}(IR\mu + R\tau_i + IR\beta_j) + \frac{1}{IJR}(\mu + JR \sum_i \tau_i IR \sum_j + \beta_j) \\ &= \mu + \tau_i + \beta_j - \mu - \tau_i - \frac{1}{J} \sum_j \beta_j - \mu - \frac{1}{I} \sum_i \tau_i - \beta_j + \mu + \frac{1}{I} \sum_i \tau_i + \frac{1}{J} \sum_j \beta_j. \end{aligned}$$

Portanto,

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

Além disso,

$$\begin{aligned}
\text{Var}(\hat{\epsilon}_{ij}) &= E[\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}]^2 \\
&= E\left[\frac{1}{R} \sum_r y_{ijr} - \frac{1}{JR} \sum_{jr} y_{ijr} - \frac{1}{IR} \sum_{ir} y_{ijr} + \frac{1}{IJR} \sum_{ijr} y_{ijr}\right]^2 \\
&= E\left[\underbrace{\frac{1}{R^2} \left(\sum_r y_{ijr}\right)^2}_{(1)} + \underbrace{\frac{1}{J^2 R^2} \left(\sum_{jr} y_{ijr}\right)^2}_{(2)} + \underbrace{\frac{1}{I^2 R^2} \left(\sum_{ir} y_{ijr}\right)^2}_{(3)}\right. \\
&\quad + \underbrace{\frac{1}{I^2 J^2 R^2} \left(\sum_{ijr} y_{ijr}\right)^2}_{(4)} - \underbrace{\frac{2}{JR^2} \left(\sum_r y_{ijr}\right) \left(\sum_{jr} y_{ijr}\right)}_{(5)} - \underbrace{\frac{2}{IR^2} \left(\sum_r y_{ijr}\right) \left(\sum_{ir} y_{ijr}\right)}_{(6)} \\
&\quad + \underbrace{\frac{2}{IJR^2} \left(\sum_r y_{ijr}\right) \left(\sum_{ijr} y_{ijr}\right)}_{(7)} + \underbrace{\frac{2}{IJR^2} \left(\sum_{jr} y_{ijr}\right) \left(\sum_{ir} y_{ijr}\right)}_{(8)} \\
&\quad \left. - \underbrace{\frac{2}{IJ^2 R^2} \left(\sum_{jr} y_{ijr}\right) \left(\sum_{ijr} y_{ijr}\right)}_{(9)} - \underbrace{\frac{2}{I^2 JR^2} \left(\sum_{ir} y_{ijr}\right) \left(\sum_{ijr} y_{ijr}\right)}_{(10)}\right]
\end{aligned}$$

onde,

$$\begin{aligned}
(1) &= \frac{1}{R^2} E(R\mu + R\tau_i + R\beta_j + R\epsilon_{ij} + \sum_r \epsilon_{ijr})^2 \\
&= \frac{E}{R^2} (R^2\mu^2 + R^2\tau_i^2 + R^2\beta_j^2 + R^2\epsilon_{ij}^2 + \left(\sum_r \epsilon_{ijr}\right)^2 \\
&\quad + 2R^2\mu\tau_i + 2R^2\mu\beta_j + 2R^2\tau_i\beta_j + dp) \\
&= \mu^2 + \tau_i^2 + \beta_j^2 + \sigma_\epsilon^2 + \frac{1}{R}\sigma^2 + 2\mu\tau_i + 2\mu\beta_j + 2\tau_i\beta_j \\
(2) &= \frac{1}{J^2 R^2} E(JR\mu + JR\tau_i + R \sum_j \beta_j + R \sum_j \epsilon_{ij} + \sum_{jr} \epsilon_{ijr})^2 \\
&= \frac{E}{J^2 R^2} (J^2 R^2 \mu^2 + J^2 R^2 \tau_i^2 + R^2 \left(\sum_j \epsilon_{ij}\right)^2 + \left(\sum_{jr} \epsilon_{ijr}\right)^2 + 2J^2 R^2 \mu\tau_i + dp) \\
&= \mu^2 + \tau_i^2 + \frac{1}{J}\sigma_\epsilon^2 + \frac{1}{JR}\sigma^2 + 2\mu\tau_i \\
(3) &= \frac{1}{I^2 R^2} E(IR\mu + R \sum_i \tau_i + IR\beta_j + R \sum_i \epsilon_{ij} + \sum_{ir} \epsilon_{ijr})^2
\end{aligned}$$

$$\begin{aligned}
&= \frac{E}{I^2 R^2} [I^2 R^2 \mu^2 + I^2 R^2 \beta_j^2 + R^2 \left(\sum_i \epsilon_{ij} \right)^2 + \left(\sum_{ir} \epsilon_{ijr} \right)^2 + 2I^2 R^2 \mu \beta_j] \\
&= \mu^2 + \beta_j^2 \frac{1}{I} \sigma_\epsilon^2 + \frac{1}{IR} \sigma^2 + 2\mu \beta_j \\
(4) \quad &= \frac{1}{I^2 J^2 R^2} E [I J R \mu + J R \sum_i \tau_i + I R R \sum_j \beta_j + R \sum_{ij} \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr}]^2 \\
&= \frac{E}{I^2 J^2 R^2} [I^2 J^2 R^2 \mu^2 + R^2 \left(\sum_{ij} \epsilon_{ij} \right)^2 + \left(\sum_{ijr} \epsilon_{ijr} \right)^2 + dp] \\
&= \mu^2 + \frac{1}{IJ} \sigma_\epsilon^2 + \frac{1}{IJR} \sigma^2 \\
(5) \quad &= \frac{1}{JR^2} E [(R\mu + R\tau_i + R\beta_j + R\epsilon_{ij} + \sum_r \epsilon_{ijr}) (JR\mu + JR\tau_i + R \sum_j \beta_j \\
&\quad + R \sum_j \epsilon_{ij} + \sum_{jr} \epsilon_{ijr})] \\
&= \frac{1}{JR^2} E [JR^2 \mu^2 + JR^2 \mu \tau_i + JR^2 \mu \tau_i + JR^2 \tau_i^2 + JR^2 \mu \beta_j \\
&\quad + JR^2 \tau_i \beta_j + R^2 \epsilon_{ij} \left(\sum_j \epsilon_{ij} \right) + \left(\sum_r \epsilon_{ijr} \right) \left(\sum_{jr} \epsilon_{ijr} \right) + dp] \\
&= \frac{1}{JR^2} (JR^2 \mu^2 + 2JR^2 \mu \tau_i + JR^2 \tau_i^2 + JR^2 \mu \beta_j + JR^2 \tau_i \beta_j + R^2 \sigma_\epsilon^2 + R\sigma^2) \\
&= \mu^2 + \tau_i^2 + 2\mu \tau_i + \mu \beta_j + \tau_i \beta_j + \frac{1}{J} \sigma_\epsilon^2 + \frac{1}{JR} \sigma^2 \\
(6) \quad &= \frac{1}{IR^2} E [(R\mu + R\tau_i + R\beta_j + R\epsilon_{ij} + \sum_r \epsilon_{ijr}) (IR\mu + R \sum_i \tau_i \\
&\quad + IR\beta_j + R \sum_i \epsilon_{ij} + \sum_{ir} \epsilon_{ijr})] \\
&= \frac{1}{IR^2} E [IR^2 \mu^2 + IR^2 \mu \beta_j + IR^2 \mu \tau_i + IR^2 \tau_i \beta_j + IR^2 \mu \beta_j \\
&\quad + IR^2 \beta_j^2 + R^2 \epsilon_{ij} \left(\sum_i \epsilon_{ij} \right) + \left(\sum_r \epsilon_{ijr} \right) \left(\sum_{ir} \epsilon_{ijr} \right) + dp] \\
&= \frac{1}{IR^2} (IR^2 \mu^2 + 2IR^2 \mu \beta_j + IR^2 \mu \tau_i + IR^2 \tau_i \beta_j + IR^2 \beta_j^2 + R^2 \sigma_\epsilon^2 + R\sigma^2) \\
&= \mu^2 + \beta_j^2 + 2\mu \beta_j + \mu \tau_i + \tau_i \beta_j + \frac{1}{I} \sigma_\epsilon^2 + \frac{1}{IR} \sigma^2 \\
(7) \quad &= \frac{1}{I J R^2} E [(R\mu + R\tau_i + R\beta_j + R\epsilon_{ij} + \sum_r \epsilon_{ijr}) (I J R \mu + J R \sum_i \tau_i \\
&\quad + I R \beta_j + R \sum_i \epsilon_{ij} + \sum_{ir} \epsilon_{ijr})]
\end{aligned}$$

$$\begin{aligned}
& + IRR \sum_j \beta_j + R \sum_{ij} \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr}] \\
& = \frac{1}{IJR^2} E[IJR^2 \mu^2 + IJR^2 \mu \tau_i + IJR^2 \mu \beta_j + R^2 \epsilon_{ij} \left(\sum_{ij} \epsilon_{ij} \right) \\
& \quad + \left(\sum_r \epsilon_{ijr} \right) \left(\sum_{ijr} \epsilon_{ijr} \right) + dp] \\
& = \frac{1}{IJR^2} (IJR^2 \mu^2 + IJR^2 \mu \tau_i + IJR^2 \mu \beta_j + R^2 \sigma_\epsilon^2 + \sigma^2) \\
& = \mu^2 + \mu \tau_i + \mu \beta_j + \frac{1}{IJ} \sigma_\epsilon^2 + \frac{1}{IJR} \sigma^2 \\
(8) \quad & = \frac{1}{IJR^2} E[(JR\mu + JR\tau_i + R \sum_j \beta_j + R \sum_j \epsilon_{ij} + \sum_{jir} \epsilon_{ijr})(IR\mu + R \sum_i \tau_i \\
& \quad + IR\beta_j + R \sum_i \epsilon_{ij} + \sum_{ir} \epsilon_{ijr})] \\
& = \frac{1}{IJR^2} E[IJR^2 \mu^2 + IJR^2 \mu \beta_j + IJR^2 \mu \tau_i + IJR^2 \tau_i \beta_j + R^2 \left(\sum_j \epsilon_{ij} \right) \left(\sum_j \epsilon_{ij} \right) \\
& \quad + \left(\sum_{jr} \epsilon_{ijr} \right) \left(\sum_{ir} \epsilon_{ijr} \right) + dp] \\
& = \frac{1}{IJR^2} (IJR^2 \mu^2 + IJR^2 \mu \tau_i + IJR^2 \mu \beta_j + IJR^2 \tau_i \beta_j + R^2 \sigma_\epsilon^2 + \sigma^2) \\
& = \mu^2 + \mu \tau_i + \mu \beta_j + \tau_i \beta_j + \frac{1}{IJ} \sigma_\epsilon^2 + \frac{1}{IJR} \sigma^2 \\
(9) \quad & = \frac{1}{IJ^2 R^2} E[(JR\mu + JR\tau_i + R \sum_j \beta_j + R \sum_j \epsilon_{ij} + \sum_{jir} \epsilon_{ijr})(IJR\mu \\
& \quad + JR \sum_i \tau_i + IR \sum_j \beta_j + R \sum_{ij} \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr})] \\
& = \frac{1}{IJ^2 R^2} E[IJ^2 R^2 \mu^2 + IJ^2 R^2 \mu \tau_i + R^2 \left(\sum_j \epsilon_{ij} \right) \left(\sum_{ij} \epsilon_{ij} \right) \\
& \quad + \left(\sum_{jr} \epsilon_{ijr} \right) \left(\sum_{ijr} \epsilon_{ijr} \right) + dp] \\
& = \frac{1}{IJ^2 R^2} (IJ^2 R^2 \mu^2 + IJ^2 R^2 \mu \tau_i + JR^2 \sigma_\epsilon^2 + JR\sigma^2) \\
& = \mu^2 + \mu \tau_i + \frac{1}{IJ} \sigma_\epsilon^2 + \frac{1}{IJR} \sigma^2 \\
(10) \quad & = \frac{1}{I^2 JR^2} E[(IR\mu + R \sum_i \tau_i + IR\beta_j + R \sum_i \epsilon_{ij} + \sum_{ir} \epsilon_{ijr})(IJR\mu
\end{aligned}$$

$$\begin{aligned}
& +JR \sum_i \tau_i + IR \sum_j \beta_j + R \sum_{ij} \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr} \\
& = \frac{1}{I^2 JR^2} E[I^2 JR^2 \mu^2 + I^2 JR^2 \mu \beta_j + R^2 \left(\sum_j \epsilon_{ij} \right) \left(\sum_{ij} \epsilon_{ij} \right) \\
& \quad + \left(\sum_{ir} \epsilon_{ijr} \right) \left(\sum_{ijr} \epsilon_{ijr} \right) + dp] \\
& = \frac{1}{I^2 JR^2} (I^2 JR^2 \mu^2 + I^2 JR^2 \mu \beta_j + IR^2 \sigma_\epsilon^2 + IR \sigma^2) \\
& = \mu^2 + \mu \beta_j + \frac{1}{IJ} \sigma_\epsilon^2 + \frac{1}{IJR} \sigma^2
\end{aligned}$$

Agora,

$$\begin{aligned}
Var(\hat{\epsilon}_{ij}) & = \mu^2 + \tau_i^2 + \beta_j^2 + \sigma_\epsilon^2 + \frac{1}{R} \sigma^2 + 2\mu\tau_i + 2\mu\beta_j + 2\tau_i\beta_j \\
& \quad + \mu^2 + \tau_i^2 + \frac{1}{J} \sigma_\epsilon^2 + \frac{1}{JR} \sigma^2 + 2\mu\tau_i \\
& \quad + \mu^2 + \beta_j^2 + \frac{1}{I} \sigma_\epsilon^2 + \frac{1}{IR} \sigma^2 + 2\mu\beta_j + \mu^2 + \frac{1}{IJ} \sigma_\epsilon^2 + \frac{1}{IJR} \sigma^2 \\
& \quad - 2\mu^2 - 2\tau_i^2 - 4\mu\tau_i - 2\mu\beta_j - 2\tau_i\beta_j - \frac{2}{J} \sigma_\epsilon^2 - \frac{2}{JR} \sigma^2 \\
& \quad - 2\mu^2 - 2\beta_j^2 - 4\mu\beta_j - 2\mu\tau_i - 2\tau_i\beta_j - \frac{2}{I} \sigma_\epsilon^2 - \frac{2}{IR} \sigma^2 \\
& \quad - 2\mu^2 - 2\mu\tau_i - 2\mu\beta_j - \frac{2}{IJ} \sigma_\epsilon^2 - \frac{2}{IJR} \sigma^2 \\
& \quad + 2\mu^2 + 2\mu\tau_i + 2\mu\beta_j + 2\tau_i\beta_j + \frac{2}{IJ} \sigma_\epsilon^2 + \frac{2}{IJR} \sigma^2 \\
& \quad - 2\mu^2 - 2\mu\tau_i - \frac{2}{IJ} \sigma_\epsilon^2 - \frac{2}{IJR} \sigma^2 - 2\mu^2 - 2\mu\beta_j - \frac{2}{IJ} \sigma_\epsilon^2 - \frac{2}{IJR} \sigma^2 \\
Var(\hat{\epsilon}_{ij}) & = 0\mu^2 + 0\tau_i^2 + 0\beta_j^2 + 0\mu\tau_i + 0\mu\beta_j + 0\tau_i\beta_j \\
& \quad + \left(1 + \frac{1}{J} + \frac{1}{I} + \frac{1}{IJ} - \frac{2}{J} - \frac{2}{I} - \frac{2}{IJ} + \frac{2}{IJ} - \frac{2}{IJ} - \frac{2}{IJ} \right) \sigma_\epsilon^2 \\
& \quad + \left(\frac{1}{R} + \frac{1}{JR} + \frac{1}{IR} + \frac{1}{IJR} - \frac{2}{JR} - \frac{2}{IR} + \frac{4}{IJR} - \frac{4}{IJR} \right) \sigma^2 \\
Var(\hat{\epsilon}_{ij}) & = \left(1 - \frac{1}{J} - \frac{1}{I} + \frac{1}{IJ} \right) \sigma_\epsilon^2 + \left(\frac{1}{R} - \frac{1}{JR} - \frac{1}{IR} + \frac{1}{IJR} \right) \sigma^2 \\
& = \frac{1}{IJ} (IJ - I - J + 1) \sigma_\epsilon^2 + \frac{1}{IJR} (IJ - I - J + 1) \sigma^2 \\
& = \frac{1}{IJ} (J(I-1) - (I-1)) \sigma_\epsilon^2 + \frac{1}{IJR} (J(I-1) - (I-1)) \sigma^2 \\
& = \frac{(I-1)(J-1)}{IJ} \sigma_\epsilon^2 + \frac{(I-1)(J-1)}{IJR} \sigma^2 \\
Var(\hat{\epsilon}_{ij}) & = \frac{(I-1)(J-1)}{IJR} (R\sigma_\epsilon^2 + \sigma^2).
\end{aligned}$$

Assim sendo, conclui-se que

$$\hat{\epsilon}_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...} \sim N\left(0; \frac{(I-1)(J-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)\right), \quad (2.35)$$

ou seja, o estimador do erro entre, $\hat{\epsilon}_{ij}$, é distribuído como uma normal de média zero e variância $\frac{(I-1)(J-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)$.

2.6.11 Distribuição de probabilidade da $SQ_{\text{Erro Entre}}$

Partindo da expressão (2.35) e lembrando dos resultados da teoria de probabilidade para a distribuição normal padrão, vem

$$\frac{\hat{\epsilon}_{ij} - E(\hat{\epsilon}_{ij})}{\sqrt{\frac{(I-1)(J-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)}} \sim N(0; 1)$$

ou

$$\begin{aligned} \frac{\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...} - 0}{\sqrt{\frac{(I-1)(J-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)}} &\sim N(0; 1) \Rightarrow \frac{(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2}{\frac{(I-1)(J-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)} \sim \chi_{(1)}^2 \Rightarrow \\ \frac{(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2}{\sigma^2 + R\sigma_\epsilon^2} &= \frac{(I-1)(J-1)}{IJR} \chi_{(1)}^2 \Rightarrow \\ \frac{\sum_{ijr} (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2}{\sigma^2 + R\sigma_\epsilon^2} &= \frac{SQ_{\text{Erro Entre}}}{\sigma^2 + R\sigma_\epsilon^2} = \frac{IJR(I-1)(J-1)}{IJR} \chi_{(1)}^2. \end{aligned}$$

Portanto,

$$\frac{SQ_{\text{Erro Entre}}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{[(I-1)(J-1)]}^2. \quad (2.36)$$

Em palavras, a soma de quadrados do erro entre parcelas dividida por $\sigma^2 + R\sigma_\epsilon^2$ tem uma distribuição de qui-quadrado com $(I-1)(J-1)$ graus de liberdade.

2.6.12 Distribuição de probabilidade de $\hat{\epsilon}_{ijr}$, o estimador de ϵ_{ijr} .

Finalmente, as características da distribuição do estimador do erro dentro, $\hat{\epsilon}_{ijr}$, é obtido a partir da expressão (2.10), isto é,

$$\hat{\epsilon}_{ijr} = y_{ijr} - \hat{y}_{ijr} = (y_{ijr} - \bar{y}_{...}) - (\bar{y}_{i..} - \bar{y}_{...}) - (\bar{y}_{.j.} - \bar{y}_{...}) - (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})$$

na qual pode-se observar que $\hat{\varepsilon}_{ijr}$ é uma combinação linear dos y_{ijr} 's os quais seguem distribuição normal. Portanto, $\hat{\varepsilon}_{ijr}$ também segue uma distribuição normal. Isto posto, é necessário, agora, saber quais as características da distribuição de $\hat{\varepsilon}_{ijr}$.

$$\begin{aligned}
 E(\hat{\varepsilon}_{ijr}) &= E \left[y_{ijr} - \frac{1}{R} \sum_r y_{ijr} \right] \\
 &= E \left[\mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr} \right. \\
 &\quad \left. - \frac{1}{R} \left(R\mu + R\tau_i + R\beta_j + R\epsilon_{ij} + \sum_r \varepsilon_{ijr} \right) \right] \\
 &= E \left[\mu + \tau_i + \beta_j + \epsilon_{ij} + \varepsilon_{ijr} - \mu - \tau_i - \beta_j - \epsilon_{ij} - \frac{1}{R} \sum_r \varepsilon_{ijr} \right].
 \end{aligned}$$

Logo,

$$E(\hat{\varepsilon}_{ijr}) = E \left[\varepsilon_{ijr} - \frac{1}{R} \sum_r \varepsilon_{ijr} \right] = 0.$$

Além do mais,

$$\begin{aligned}
 Var(\hat{\varepsilon}_{ijr}) &= [\hat{\varepsilon}_{ijr} - E(\hat{\varepsilon}_{ijr})]^2 = E \left[\varepsilon_{ijr} - \frac{1}{R} \sum_r \varepsilon_{ijr} \right]^2 \\
 &= E \left[\varepsilon_{ijr}^2 + \frac{1}{R^2} \left(\sum_r \varepsilon_{ijr} \right)^2 - \frac{2}{R} \varepsilon_{ijr} \sum_r \varepsilon_{ijr} \right] \\
 &= \sigma^2 + \frac{1}{R^2} E(\varepsilon_{ij1}^2 + \varepsilon_{ij2}^2 + \dots + \varepsilon_{ijR}^2 + dp) \\
 &\quad - \frac{2}{R} E[\varepsilon_{ijr}(\varepsilon_{ij1} + \varepsilon_{ij2} + \dots + \varepsilon_{ijr} + \dots + \varepsilon_{ijR})] \\
 &= \sigma^2 + \frac{R}{R^2} \sigma^2 - \frac{2}{R} \sigma^2 = \sigma^2 - \frac{\sigma^2}{R} = \frac{1}{R} (R-1) \sigma^2 \\
 Var(\hat{\varepsilon}_{ijr}) &= \frac{(R-1)}{R} \sigma^2.
 \end{aligned}$$

Portanto,

$$\hat{\varepsilon}_{ijr} = (y_{ijr} - \hat{y}_{ijr}) = y_{ijr} - \bar{y}_{ij} \sim N \left(0; \frac{(R-1)}{R} \sigma^2 \right).$$

Isto é, o estimador do erro dentro, $\hat{\varepsilon}_{ijr}$, segue uma distribuição normal com média *zero* e variância $\frac{(R-1)}{R} \sigma^2$.

Além disso,

$$\begin{aligned} \frac{y_{ijr} - \bar{y}_{ij.} - 0}{\sqrt{\frac{(R-1)}{R}\sigma^2}} &\sim N(0; 1) \Rightarrow \frac{(y_{ijr} - \bar{y}_{ij.})^2}{\frac{(R-1)}{R}\sigma^2} \sim \chi_{(1)}^2 \Rightarrow \\ \frac{(y_{ijr} - \bar{y}_{ij.})^2}{\sigma^2} &\sim \frac{R-1}{R}\chi_{(1)}^2 \Rightarrow \frac{\sum_{ijr} (y_{ijr} - \bar{y}_{ij.})^2}{\sigma^2} \sim \frac{IJR(R-1)}{R}\chi_{(1)}^2. \end{aligned}$$

Portanto,

$$\frac{\sum_{ijr} (y_{ijr} - \bar{y}_{ij.})^2}{\sigma^2} = \frac{SQ_{Erro Dentro}}{\sigma^2} \sim \chi_{[IJ(R-1)]}^2. \quad (2.37)$$

2.7 Valores Esperados das Somas de Quadrados

Os valores esperados dos quadrados médios são argumentos importantes para a compreensão da escolha das estatísticas de teste utilizadas na contrastação das hipóteses de interesse. Nesta seção serão abordados detalhadamente como estes resultados são obtidos.

Viu-se, na seção 2.5, que a soma de quadrados total é decomposta em partes componentes cujas expressões são dadas a seguir.

$$\begin{aligned} SQ_{Total} &= \sum_{ijr} y_{ijr}^2 - C, \quad \text{em que } C = \frac{y_{...}^2}{IJR} \\ SQ_{Trat} &= JR \sum_i (\bar{y}_{i..} - \bar{y}_{...})^2 = \frac{1}{JR} \sum_i y_{i..}^2 - C \\ SQ_{Blocos} &= IR \sum_j (\bar{y}_{.j.} - \bar{y}_{...})^2 = \frac{1}{IR} \sum_j y_{.j.}^2 - C \\ SQ_{Parcelas} &= R \sum_{ij} (\bar{y}_{ij.} - \bar{y}_{...})^2 = \frac{1}{R} \sum_{ij} y_{ij.}^2 - C \\ SQ_{Erro Entre} &= SQ_{Parcelas} - SQ_{Trat} - SQ_{Blocos} \\ SQ_{Erro Dentro} &= SQ_{Total} - SQ_{Parcelas} \end{aligned}$$

Os valores esperados dos quadrados médios são obtidos a partir das expressões acima.

Esperança da Soma de Quadrados Total

Em primeiro lugar optou-se pelo cálculo do valor esperado da soma de quadrados

total. Isto é,

$$E(SQ_{Total}) = E\left(\sum_{ijr} y_{ijr}^2 - C\right) = E\left(\sum_{ijr} y_{ijr}^2\right) - E(C)$$

Mas,

$$\begin{aligned} E(C) &= E\left(\frac{y^2}{IJR}\right) = \frac{1}{IJR} E\left(\sum_{ijr} y_{ijr}\right)^2 \\ &= \frac{1}{IJR} E\left(IJR\mu + JR\sum_i \tau_i + IRR\sum_j \beta_j + R\sum_{ij} \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr}\right)^2 \\ &= \frac{1}{IJR} E\left(IJR\mu + R\sum_{ij} \epsilon_{ij} + \sum_{ijr} \epsilon_{ijr}\right)^2 \\ &= \frac{1}{IJR} E\left[I^2 J^2 R^2 \mu^2 + R^2 \left(\sum_{ij} \epsilon_{ij}\right)^2 + \left(\sum_{ijr} \epsilon_{ijr}\right)^2 + dp\right] \\ &= \frac{1}{IJR} (I^2 J^2 R^2 \mu^2 + IJR^2 \sigma_\epsilon^2 + IJR\sigma^2). \end{aligned}$$

Portanto,

$$E(C) = IJR\mu^2 + R\sigma_\epsilon^2 + \sigma^2 \quad (I)$$

e

$$\begin{aligned} E\left(\sum_{ijr} y_{ijr}^2\right) &= E\left[\sum_{ijr} (\mu + \tau_i + \beta_j + \epsilon_{ij} + \epsilon_{ijr})^2\right] \\ &= E\left[\sum_{ijr} (\mu^2 + \tau_i^2 + \beta_j^2 + \epsilon_{ij}^2 + \epsilon_{ijr}^2 + 2\mu\tau_i + 2\mu\beta_j + 2\tau_i\beta_j + dp)\right] \\ &= \sum_{ijr} E(\mu^2 + \tau_i^2 + \beta_j^2 + \epsilon_{ij}^2 + \epsilon_{ijr}^2 + 2\mu\tau_i + 2\mu\beta_j + 2\tau_i\beta_j + dp) \\ &= \sum_{ijr} (\mu^2 + \tau_i^2 + \beta_j^2 + \sigma_\epsilon^2 + \sigma^2 + 2\mu\tau_i + 2\mu\beta_j + 2\tau_i\beta_j) \\ &= IJR\mu^2 + JR\sum_i \tau_i^2 + IR\sum_j \beta_j^2 + IJR\sigma_\epsilon^2 + IJR\sigma^2 \\ &\quad + 2JR\mu\sum_i \tau_i + 2IR\mu\sum_j \beta_j + 2R\sum_i \tau_i\sum_j \beta_j \\ &= IJR\mu^2 + JR\sum_i \tau_i^2 + IR\sum_j \beta_j^2 + IJR\sigma_\epsilon^2 + IJR\sigma^2 \quad (II) \end{aligned}$$

Subtraindo-se (I) de (II), vem

$$E(SQ_{Total}) = IJR\mu^2 + JR\sum_i \tau_i^2 + IR\sum_j \beta_j^2 + IJR\sigma_\epsilon^2 + IJR\sigma^2$$

$$-IJR\mu^2 - R\sigma_\epsilon^2 - \sigma^2$$

Portanto, o valor esperado da soma de quadrados total é definido por:

$$E(SQ_{Total}) = R(IJ - 1)\sigma_\epsilon^2 + (IJR - 1)\sigma^2 + JR \sum_i \tau_i^2 + IR \sum_j \beta_j^2. \quad (2.38)$$

Esperança da Soma de Quadrados de Tratamento

Em segundo lugar procurou-se deduzir o valor esperado da soma de quadrados de tratamentos, ou seja

$$E(SQ_{Trat}) = \left[\frac{1}{JR} \sum_i y_{i..}^2 - C \right] = \frac{1}{JR} E \left(\sum_i y_{i..}^2 \right) - E(C)$$

Mas,

$$\begin{aligned} \frac{1}{JR} E \left(\sum_i y_{i..}^2 \right) &= \frac{1}{JR} E \left[\sum_i \left(JR\mu + JR\tau_i + R \sum_j \beta_j + R \sum_j \epsilon_{ij} + \sum_{jr} \epsilon_{ij} \right)^2 \right] \\ &= \frac{1}{JR} E \left[\sum_i \left(J^2 R^2 \mu^2 + J^2 R^2 \tau_i^2 + R^2 \left(\sum_j \epsilon_{ij} \right)^2 + \left(\sum_{jr} \epsilon_{ir} \right)^2 \right. \right. \\ &\quad \left. \left. + 2J^2 R^2 \mu \tau_i + 0dp \right) \right] \\ &= \frac{1}{JR} \sum_i [J^2 R^2 \mu^2 + J^2 R^2 \tau_i^2 + JR^2 \sigma_\epsilon^2 + JR\sigma^2 + 2J^2 R^2 \mu \tau_i] \\ &= \frac{1}{JR} [IJ^2 R^2 \mu^2 + J^2 R^2 \sum_i \tau_i^2 + IJR^2 \sigma_\epsilon^2 + IJR\sigma^2 + 0]. \end{aligned}$$

Assim sendo, tem-se

$$\frac{1}{JR} E \left(\sum_i y_{i..}^2 \right) = IJR\mu^2 + JR \sum_i \tau_i^2 + IR\sigma_\epsilon^2 + I\sigma^2. \quad (III)$$

Subtraindo-se (I) de (III), obtém-se o valor esperado da soma de quadrados de tratamentos, isto é,

$$E(SQ_{Trat}) = (I - 1)(\sigma^2 + R\sigma_\epsilon^2) + JR \sum_i \tau_i^2. \quad (2.39)$$

A esperança do quadrado médio de tratamentos é dada pelo valor esperado da soma de quadrados de tratamentos dividida pelos seus respectivos graus de liberdade, ou seja,

$$E(QM_{Trat}) = E \left[\frac{SQ_{Trat}}{I - 1} \right] = \sigma^2 + R\sigma_\epsilon^2 + \frac{JR}{I - 1} \sum_i \tau_i^2 \quad (2.40)$$

ou

$$E(QM_{Trat}) = \sigma^2 + R\sigma_\epsilon^2 + \frac{JR}{I-1} \sum_i (\mu_i - \mu)^2, \quad (2.41)$$

em que $\mu_i = \mu + \tau_i$.

Esperança da soma de quadrados de uma combinação linear das médias dos tratamentos: h -ésimo contraste de médias dos tratamentos

De acordo com a expressão (2.34), a soma de quadrados de um contraste h é dada por

$$SQ(\hat{\Psi}_h) = \frac{\left(\sum_{i=1}^I c_{hi} y_{i..} \right)^2}{JRK_h}.$$

Desenvolvendo-se algebricamente a expressão acima, vem

$$\begin{aligned} SQ(\hat{\Psi}_h) &= \frac{\left(\sum_{i=1}^I c_{hi} y_{i..} \right)^2}{JRK_h} \\ &= \frac{1}{JRK_h} \left[\sum_{i=1}^I c_{hi} \left(JR\mu_i + R \sum_{j=1}^J \beta_j + R \sum_{j=1}^J \epsilon_{ij} + \sum_{j=1}^J \sum_{r=1}^R \epsilon_{ijr} \right) \right]^2 \\ &= \frac{1}{JRK_h} \left[J^2 R^2 \left(\sum_{i=1}^I c_{hi} \mu_i \right)^2 + R^2 \left(\sum_{i=1}^I c_{hi} \sum_{j=1}^J \epsilon_{ij} \right)^2 + \left(\sum_{i=1}^I c_{hi} \sum_{j=1}^J \sum_{r=1}^R \epsilon_{ijr} \right)^2 + dp \right]. \end{aligned}$$

Agora, calculando o valor esperado, obtém-se após algumas operações algébricas o seguinte resultado

$$E[SQ(\hat{\Psi}_h)] = E[QM(\hat{\Psi}_h)] = \sigma^2 + R\sigma_\epsilon^2 + \frac{JR}{K_h} \left(\sum_{i=1}^I c_{hi} \mu_i \right)^2. \quad (2.42)$$

uma vez que, por (2.34), há apenas um grau de liberdade associado a soma de quadrados de um contraste entre as médias dos tratamentos.

Esperança da Soma de Quadrados de Blocos

Para calcular a esperança da soma de quadrados de blocos, procedeu-se da seguinte forma,

$$\begin{aligned} E(SQ_{Blocos}) &= E \left[IR \sum_j (\bar{y}_{.j} - \bar{y}_{...})^2 \right] \\ &= E \left[\frac{1}{IR} \sum_j y_{.j}^2 - C \right] = \frac{1}{IR} E \left(\sum_j y_{.j}^2 \right) - E(C) \end{aligned}$$

Mas,

$$\begin{aligned}
\frac{1}{IR}E\left(\sum_j y_{.j}^2\right) &= \frac{1}{IR}E\left[\sum_j \left(IR\mu + R\sum_i \tau_i + IR\beta_j + R\sum_i \epsilon_{ij} + \sum_{ir} \varepsilon_{ijr} \right)^2 \right] \\
&= \frac{1}{IR}E\left[\sum_j \left(I^2R^2\mu^2 + I^2R^2\beta_j^2 + R^2\left(\sum_i \epsilon_{ij}\right)^2 + \left(\sum_{ir} \varepsilon_{ijr}\right)^2 \right. \right. \\
&\quad \left. \left. + 2I^2R^2\mu\beta_j + 0dp \right) \right] \\
&= \frac{1}{IR}\sum_j (I^2R^2\mu^2 + I^2R^2\beta_j^2 + IR^2\sigma_\epsilon^2 + IR\sigma^2 + 2I^2R^2\mu\beta_j) \\
&= \frac{1}{IR}\left(I^2JR^2\mu^2 + I^2R^2\sum_j \beta_j^2 + IJR^2\sigma_\epsilon^2 + IJR\sigma^2 + 2I^2JR^2\mu\sum_j \beta_j \right).
\end{aligned}$$

Portanto,

$$\frac{1}{IR}E\left(\sum_j y_{.j}^2\right) = IJR\mu^2 + IR\sum_j \beta_j^2 + JR^2\sigma_\epsilon^2 + J\sigma^2 \quad (IV)$$

Subtraindo-se (I) de (IV), encontrou-se,

$$\begin{aligned}
E(SQ_{Blocos}) &= IJR\mu^2 + IR\sum_j \beta_j^2 + JR^2\sigma_\epsilon^2 + J\sigma^2 - IJR\mu^2 - R\sigma_\epsilon^2 - \sigma^2 \\
&= (J-1)\sigma^2 + R(J-1)\sigma_\epsilon^2 + IR\sum_j \beta_j^2
\end{aligned}$$

$$E(SQ_{Blocos}) = (J-1)(\sigma^2 + R\sigma_\epsilon) + IR\sum_j \beta_j^2. \quad (2.43)$$

A esperança do quadrado médio de blocos é dada pelo valor esperado da soma de quadrados de blocos dividida pelos seus respectivos graus de liberdade, ou seja,

$$E(QM_{Blocos}) = E\left[\frac{SQ_{Blocos}}{J-1}\right] = \sigma^2 + R\sigma_\epsilon^2 + IR\sum_j \beta_j^2 \quad (2.44)$$

ou

$$E(QM_{Blocos}) = \sigma^2 + R\sigma_\epsilon + \frac{IR}{J-1}\sum_j (\mu_j - \mu)^2, \quad (2.45)$$

em que $\mu_j = \mu + \beta_j$.

Esperança da Soma de Quadrados de Parcelas

Para a soma de quadrados de parcelas procedeu-se como segue,

$$\begin{aligned} E(SQ_{Parcelas}) &= E \left[R \sum_{ij} (\bar{y}_{ij} - \bar{y}_{...})^2 \right] \\ &= E \left[\frac{1}{R} \sum_{ij} y_{ij}^2 - C \right] = \frac{1}{R} E \left(\sum_{ij} y_{ij}^2 \right) - E(C) \end{aligned}$$

Mas,

$$\begin{aligned} \frac{1}{R} E \left(\sum_{ij} y_{ij}^2 \right) &= \frac{1}{R} E \left[\sum_{ij} \left(R\mu + R\tau_i + R\beta_j + R\epsilon_{ij} + \sum_r \epsilon_{ijr} \right)^2 \right] \\ &= \frac{1}{R} E \left[\sum_{ij} \left(R^2\mu^2 + R^2\tau_i^2 + R^2\beta_j^2 + R^2\epsilon_{ij}^2 + \left(\sum_r \epsilon_{ijr} \right)^2 + 2R^2\mu\tau_i \right. \right. \\ &\quad \left. \left. + 2R^2\mu\beta_j + 2R^2\tau_i\beta_j + 0dp \right) \right] \\ &= \frac{1}{R} E \left(R^2\mu^2 + R^2\tau_i^2 + R^2\beta_j^2 + R^2\sigma_\epsilon^2 + \sigma^2 + 2R^2\mu\tau_i + 2R^2\mu\beta_j \right. \\ &\quad \left. + 2R^2\tau_i\beta_j \right) \\ &= \frac{1}{R} \left(IJR^2\mu^2 + JR^2 \sum_i \tau_i^2 + IR^2 \sum_j \beta_j^2 + IJR^2\sigma_\epsilon^2 + IJR\sigma^2 \right. \\ &\quad \left. + 2JR^2\mu \sum_i \tau_i + 2IR^2\mu \sum_j \beta_j + 2R^2 \sum_i \tau_i \sum_j \beta_j \right). \end{aligned}$$

Logo,

$$\frac{1}{R} E \left(\sum_{ij} y_{ij}^2 \right) = IJR\mu^2 + JR \sum_i \tau_i^2 + IR \sum_j \beta_j^2 + IJR^2\sigma_\epsilon^2 + IJ\sigma^2. \quad (V)$$

Subtraindo-se (I) de (V), obtém-se

$$\begin{aligned} E(SQ_{Parcelas}) &= IJR\mu^2 + JR \sum_i \tau_i^2 + IR \sum_j \beta_j^2 + IJR^2\sigma_\epsilon^2 + IJ\sigma^2 \\ &\quad - IJR\mu^2 - R\sigma_\epsilon^2 - \sigma^2 \\ &= (IJ - 1)\sigma^2 + R(IJ - 1)\sigma_\epsilon^2 + JR \sum_i \tau_i^2 + IR \sum_j \beta_j^2 \end{aligned}$$

$$E(SQ_{Parcelas}) = (IJ - 1)(\sigma^2 + R\sigma_\epsilon^2) + JR \sum_i \tau_i^2 + IR \sum_j \beta_j^2. \quad (2.46)$$

De modo análogo aos casos anteriores, o valor esperado do quadrado médio de parcelas

fica,

$$E(QM_{Parcelas}) = E\left(\frac{SQ_{Parcelas}}{IJ-1}\right) = (\sigma^2 + R\sigma_\epsilon^2) + \frac{JR}{IJ-1} \sum_i \tau_i^2 + \frac{IR}{IJ-1} \sum_j \beta_j^2 \quad (2.47)$$

ou

$$E(QM_{Parcelas}) = (\sigma^2 + R\sigma_\epsilon^2) + \frac{JR}{IJ-1} \sum_i (\mu_i - \mu)^2 + \frac{IR}{IJ-1} \sum_j (\mu_j - \mu)^2, \quad (2.48)$$

em que, $\mu_i = \mu + \tau_i$ e $\mu_j = \mu + \beta_j$.

Esperança da Soma de Quadrados do Erro Entre

O valor esperado da soma de quadrados do erro entre é obtida pela subtração dos valores esperados da soma de quadrados de bloco e soma de quadrados de tratamento da soma de quadrados de parcela, ou seja,

$$\begin{aligned} SQ_{Erro\ Entre} &= E(SQ_{Parcelas} - SQ_{Trat} - SQ_{Blocos}) \\ &= E(SQ_{Parcela}) - E(SQ_{Trat}) - E(SQ_{Bloco}) \\ &= (IJ-1)(\sigma^2 + R\sigma_\epsilon^2) + JR \sum_i \tau_i^2 + IR \sum_j \beta_j^2 \\ &\quad - (I-1)(\sigma^2 + R\sigma_\epsilon^2) - JR \sum_i \tau_i^2 \\ &\quad - (J-1)(\sigma^2 + R\sigma_\epsilon^2) - IR \sum_j \beta_j^2 \\ &= (\sigma^2 + R\sigma_\epsilon^2)[IJ-1 - I + 1 - J + 1] \\ E(SQ_{Erro\ Entre}) &= (I-1)(J-1)(\sigma^2 + R\sigma_\epsilon^2) \end{aligned} \quad (2.49)$$

Como feito anteriormente, tem-se a esperança do quadrado médio do erro entre da seguinte forma,

$$E(QM_{Erro\ Entre}) = E\left(\frac{SQ_{Erro\ Entre}}{(I-1)(J-1)}\right) = \sigma^2 + R\sigma_\epsilon^2; \quad (2.50)$$

Esperança da Soma de Quadrados do Erro Dentro

Obtém-se o valor esperado da soma de quadrado de erro dentro com a seguinte subtração,

$$E(SQ_{Erro\ Dentro}) = E(SQ_{Total} - SQ_{Parcelas})$$

$$\begin{aligned}
&= E(SQ_{Total}) - E(SQ_{Parcelas}) \\
&= R(IJ - 1)\sigma_\epsilon^2 + (IJR - 1)\sigma^2 + JR \sum_i \tau_i^2 + IR \sum_j \beta_j^2 \\
&\quad - (IJ - 1)(\sigma^2 + R\sigma_\epsilon^2) - JR \sum_i \tau_i^2 - IR \sum_j \beta_j^2 \\
&= IJR\sigma_\epsilon^2 - R\sigma^2 + IJR^2 - \sigma^2 - IJ\sigma^2 + \sigma^2 - IJR\sigma_\epsilon^2 + R\sigma_\epsilon^2.
\end{aligned}$$

Portanto,

$$E(SQ_{Erro Dentro}) = IJ(R - 1)\sigma^2 \quad (2.51)$$

Dividindo-se a soma de quadrado de erro dentro pelos seus respectivos graus de liberdade tem-se o valor esperado do quadrado médio do erro dentro,

$$E(QM_{Erro Dentro}) = E\left(\frac{SQ_{Erro Dentro}}{IJ(R - 1)}\right) = \sigma^2. \quad (2.52)$$

Portanto, o valor esperado da $SQ_{Erro Dentro}$ dividido pela variância do erro dentro segue uma distribuição de qui-quadrado com $IJ(R - 1)$ graus de liberdade.

Os resultados obtidos nesta seção encontram-se apresentados de forma resumida na Tabela 4, como sugere Barbin (1993).

Tabela 4: Análise da variância com os valores esperados dos quadrados médios

F.V.	G.L.	S.Q.	E(QM)
Tratamento	I-1	$\frac{1}{JR} \sum_i y_{i..}^2 - C$	$\sigma^2 + R\sigma_\epsilon^2 + \frac{JR}{I-1} \sum_i (\mu_i - \mu)^2$
Blocos	J-1	$\frac{1}{IR} \sum_j y_{.j}^2 - C$	$\sigma^2 + R\sigma_\epsilon^2 + \frac{IR}{J-1} \sum_j (\mu_j - \mu)^2$
Erro Entre Parcelas	(I-1)(J-1)	$SQ_{Parc} - SQ_{Trat} - SQ_{Blocos}$	$\sigma^2 + R\sigma_\epsilon^2$
Parcelas	(IJ-1)	$\frac{1}{R} \sum_{ij} y_{ij}^2 - C$	-
Erro Dentro	IJ(R-1)	$SQ_{Total} - SQ_{Parcelas}$	σ^2
Total	IJR-1	$\sum_{ijr} y_{ijr}^2 - C$	-

A tabela 5 contém informações acerca dos parâmetros associados ao modelo matemático, aos seus estimadores e suas respectivas distribuições de probabilidade, bem como, as distribuições de probabilidade das somas de quadrados associadas.

Tabela 5: Características, seus estimadores e distribuições de probabilidade de estatísticas associadas.

Característica	Estimador da característica	Distribuição de Probabilidade do Estimador	Distribuição de Probabilidade das Estatísticas de Interesse sob H_0
μ	$\hat{\mu} = m = \bar{y}_{...}$	$\hat{\mu} \sim N\left[\mu; \frac{1}{IJR}(\sigma^2 + R\sigma_\epsilon^2)\right]$	Sob $H_0 : \mu = 0$, $\frac{C}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{(1)}^2$
τ_i	$\hat{\tau}_i = t_i = \bar{y}_{i..} - \bar{y}_{...}$	$\hat{\tau}_i \sim N\left[\tau_i; \frac{(I-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)\right]$	Sob $H_0 : \tau_i = 0, \forall i$, $\frac{SQ_{Trat}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{(I-1)}^2$
μ_i	$\hat{\mu}_i = m_i = \bar{y}_{i..}$	$\hat{\mu}_i \sim N\left[\mu_i; \frac{1}{JR}(\sigma^2 + R\sigma_\epsilon^2)\right]$	Sob $H_0 : \mu_i = \mu_0$, $\frac{\hat{\mu}_i - \mu_0}{\sqrt{\frac{QM_{Erro_Entre}}{JR}}} \sim t_{[(I-1)(J-1)]}$
β_j	$\hat{\beta}_j = b_j = \bar{y}_{.j} - \bar{y}_{...}$	$\hat{\beta}_j \sim N\left[\beta_j; \frac{(J-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)\right]$	Sob $H_0 : \beta_j = 0, \forall j$, $\frac{SQ_{Blocos}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{(J-1)}^2$
μ_j	$\hat{\mu}_j = m_j = \bar{y}_{.j}$	$\mu_j \sim N\left[\mu_j; \frac{1}{IR}(\sigma^2 + R\sigma_\epsilon^2)\right]$	Sob $H_0 : \mu_j = \mu_0$, $\frac{\hat{\mu}_j - \mu_0}{\sqrt{\frac{QM_{Erro_Entre}}{IR}}} \sim t_{[(I-1)(J-1)]}$
$\Psi_h = \sum_{i=1}^I c_{hi}\mu_i$	$\hat{\Psi}_h = \sum_{i=1}^I c_{hi}\bar{y}_{i..}$	$\hat{\Psi}_h \sim N\left[\sum_{i=1}^I c_{hi}\mu_i, \frac{1}{JR}(\sigma^2 + R\sigma_\epsilon^2)\right]$	Sob $H_0 : \Psi_h = 0$, $\frac{\left(\sum_{i=1}^I c_{hi}\bar{y}_{i..}\right)^2}{JRK_h(\sigma^2 + R\sigma_\epsilon^2)} \sim \chi_{(1)}^2$
ϵ_{ij}	$\hat{\epsilon}_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j} - \bar{y}_{...}$	$\hat{\epsilon}_{ij} \sim N\left[0; \frac{(I-1)(J-1)}{IJR}(\sigma^2 + R\sigma_\epsilon^2)\right]$	$\frac{SQ_{Erro_Entre}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{[(I-1)(J-1)]}^2$
ϵ_{ijr}	$\hat{\epsilon}_{ijr} = y_{ijr} - \bar{y}_{ij.}$	$\hat{\epsilon}_{ijr} \sim N\left[0; \frac{IJ(R-1)}{IJR}\sigma^2\right]$	$\frac{SQ_{Erro_Dentro}}{\sigma^2} \sim \chi_{[IJ(R-1)]}^2$
y_{ijr}	$\hat{y}_{ijr} = \bar{y}_{ij.}$	-	-

2.8 Comparações múltiplas das médias duas a duas

Como o desenvolvimento da teoria dos métodos de comparações múltiplas não faz parte dos objetivos deste trabalho, achou-se conveniente apresentar simplesmente um procedimento prático para comparar os possíveis pares de médias dos tratamentos por meio do teste de Tukey.

Para testar as hipóteses do tipo $H_0 : \mu_i = \mu_{i'}$ contra $H_1 : \mu_i \neq \mu_{i'}$, ao nível de significância α , calcula-se a *Diferença Mínima Significativa - DMS* por meio da expressão

$$DMS = q_{[I; (I-1)(J-1); \alpha]} \sqrt{\frac{QM_{Erro\ Entre}}{JR}} \quad (2.53)$$

na qual $q_{[I; (I-1)(J-1); \alpha]}$ é o valor crítico da aptitude estudentizada de Tukey para $I =$ número de tratamentos envolvidos no ensaio, $(I-1)(J-1) =$ número de graus de liberdade do *Erro Entre* e nível de significância α (em geral $\alpha = 0,05$).

Em seguida calcular os valores absolutos das diferenças entre as estimativas dos possíveis pares de médias dos tratamentos envolvidos no experimento, $|\bar{y}_i - \bar{y}_{i'}|$, $i \neq i'$, $i, i' = 1, 2, \dots, I$. Finalmente, adotar a regra de decisão: Rejeiar H_0 em favor de H_1 , ao nível de significância α se, e somente se $|\bar{y}_i - \bar{y}_{i'}| > DMS$. Este procedimento será ilustrado na seção 3.

2.9 Análises estatísticas

Tal como no experimento em blocos ao acaso usual, as análises estatísticas de um experimento em blocos casualizados com repetição do conjunto de tratamentos dentro dos blocos, em geral, levam em consideração as seguintes hipóteses: Hipóteses sobre a não existência de efeito dos tratamentos sobre a variável resposta, $H_0^{(\tau)} : \tau_i = 0, \forall i = 1, 2, \dots, I$ (ou equivalentemente, $H_0^{(\tau)} : \mu_1 = \dots = \mu_I = \mu$); Hipótese sobre a não existência do efeito dos blocos: $H_0^{(\beta)} : \beta_j = 0, \forall j = 1, 2, \dots, J$. De modo geral, o pesquisador planeja seus experimentos em blocos casualizados com o objetivo apenas de proporcionar ambientes homogêneos (blocos) dentro dos quais ele distribui de modo aleatório um conjunto de tratamentos (ou mais de um conjunto), favorecendo-se o controle local. Assim sendo, quase sempre, o pesquisador não tem interesse em fazer inferência sobre o efeito dos blocos, fixando-se apenas nas análises baseadas no efeito dos tratamentos. Neste sentido, as análises são conduzidas priorizando-se os contrastes entre as médias

dos tratamentos de interesse do pesquisador, bem como, as comparações múltiplas das médias. Na seção a seguir, serão discutidas as bases teóricas que possibilitam essa possíveis análises.

2.9.1 Hipóteses sobre tratamento

Esta hipótese pode ser representada de duas maneiras, a saber:

$$\begin{aligned}
 & a) \left\{ \begin{array}{l} H_0^{(\tau)} : \tau_i = 0, \forall i = 1, 2, \dots, I \\ vs \\ H_1^{(\tau)} : \tau_i \neq 0, \text{ para pelo menos um } \tau_i \end{array} \right. \\
 & \text{ou} \\
 & b) \left\{ \begin{array}{l} H_0^{(\tau)} : \mu_1 = \mu_2 = \dots = \mu_I = \mu \\ vs \\ H_1^{(\tau)} : \mu_i \neq \mu_{i'}, \text{ para pelo menos um par } (\mu_i, \mu_{i'}) \ i \neq i' = 1, 2, \dots, I \end{array} \right.
 \end{aligned} \tag{2.54}$$

Considerando-se os resultados obtidos no desenvolvimento da teoria, pode ser observado que:

a.1) De acordo com a equação (2.50), $E(QM_{Erro\ Entre}) = \sigma^2 + R\sigma_\epsilon^2$. Isto é, o valor esperado do quadrado médio do erro entre é um estimador não viciado para $\sigma^2 + R\sigma_\epsilon^2$ independentemente de que $H_0^{(\tau)}$ seja verdadeiro;

a.2) Pela expressão (2.41), $E(QM_{Trat}) = \sigma^2 + R\sigma_\epsilon^2 + \frac{JR}{I-1} \sum_{i=1}^I (\mu_i - \mu)^2$ e se a hipótese nula $H_0^{(\tau)} : \mu_1 = \dots = \mu_I = \mu$ é verdadeira, então $\frac{JR}{I-1} \sum_{i=1}^I (\mu_i - \mu)^2 = 0$ e QM_{Trat} será um estimador não viciado para $\sigma^2 + R\sigma_\epsilon^2$. No entanto, se $H_1^{(\tau)}$ é verdadeira $E(QM_{Trat}) > \sigma^2 + R\sigma_\epsilon^2$. Assim sendo, é razoável comparar QM_{Trat} com $QM_{Erro\ Entre}$ para se efetuar o teste da hipótese $H_0^{(\tau)}$, tendo em vista que quanto maior for o QM_{Trat} comparado com $QM_{Erro\ Entre}$ mais evidência se tem de que as médias dos tratamentos são diferentes entre si (ou que o efeito dos tratamentos não são nulos).

a.3) Conforme equação (2.20), tem-se que sob $H_0^{(\tau)}$, $\frac{SQ_{Trat}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{[I-1]}^2$ e independentemente de que $H_0^{(\tau)}$ se verifique e pela expressão (2.35), $\frac{SQ_{Erro\ Entre}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{[(I-1)(J-1)]}^2$.

Como se sabe, de acordo com a teoria de estatística matemática, [ver Rohatgi (1976), Roussas (1997), dentre outros], se uma variável aleatória U segue uma distribuição

de qui-quadrado com ν_1 graus de liberdade e uma outra variável aleatória V segue uma distribuição de qui-quadrado com ν_2 graus de liberdade e, além disso, U e V são variáveis aleatórias independentes, então a razão entre a variável aleatória U dividida pelos seus graus de liberdade e a variável aleatória V dividida pelos seus graus de liberdade, segue uma distribuição F de *Snedecor e Cochran* com ν_1 graus de liberdade do numerador e ν_2 graus de liberdade do denominador a qual pode ser escrita como

$$\frac{U/\nu_1}{V/\nu_2} \sim F_{[\nu_1, \nu_2]} \quad (2.55)$$

Tomando como base este resultado, e considerando-se as distribuições de probabilidade das somas de quadrados apresentadas na Tabela 5, pode-se verificar que

$$F = \frac{\frac{SQ_{Trat}}{\sigma^2 + R\sigma_\epsilon^2}}{(I-1)} = \frac{QM_{Trat}}{QM_{Erro Entre}} \sim F_{[(I-1), (I-1)(J-1)]} \quad (2.56)$$

que será a estatística de teste para testar a hipótese $H_0^{(\tau)} : \tau_i = 0$, para todo i , vs $H_0^{(\tau)} : \tau_i \neq 0$ para pelo menos um τ_i . Rejeita-se $H_0^{(\tau)}$, ao nível de significância α , se $F = \frac{QM_{Trat}}{QM_{Erro Entre}} > F_{[(I-1), (I-1)(J-1), \alpha]}$, em que $F_{[(I-1), (I-1)(J-1), \alpha]}$ é o $100(1 - \alpha)$ -ésimo percentil superior da distribuição F com $(I - 1)$ graus de liberdade do numerador e $(I - 1)(J - 1)$ graus de liberdade do denominador.

- a.4)** Como pode ser observado na expressão (2.50), o valor esperado do $QM_{Erro Entre}$ é $\sigma^2 + R\sigma_\epsilon^2$ independentemente de que a hipótese $H_0^{(\Psi)} : \Psi_h = 0$ se verifique ou não. Isto é, $QM_{Erro Entre}$ é um estimador não tendencioso para $\sigma^2 + R\sigma_\epsilon^2$.
- a.5)** Da expressão (2.42), se $H_0^{(\Psi)} : \sum_{i=1}^I c_{hi}\mu_i = 0$ é verdadeiro, $QM(\hat{\Psi}_h)$ também é um estimador não tendencioso para $\sigma^2 + R\sigma_\epsilon^2$. Porém, se $H_0^{(\Psi)}$ for falsa, $\sum_{i=1}^I c_{hi}\mu_i \neq 0$ e $E[QM(\hat{\Psi}_h)] > \sigma^2 + R\sigma_\epsilon^2$. Portanto, para se efetuar um teste de hipótese sobre um contraste de médias dos tratamentos é intuitivo que seja feita uma comparação entre o quadrado médio do contraste e o quadrado médio do erro entre. Pois, espera-se que quanto maior for $QM(\hat{\Psi}_h)$ comparado com $QM_{Erro Entre}$, mais evidência se tem de que o contraste $\Psi_h = \sum_{i=1}^I c_{hi}\mu_i$ não é nulo e quanto mais aproxima-se de zero mais evidência se tem a favor de H_0 .
- a.6)** Conforme equação (2.34), tem-se que sob $H_0^{(\Psi)}$, a estatística $\frac{SQ(\hat{\Psi}_h)}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{[1]}^2$ e, pela expressão (2.35), $\frac{SQ_{Erro Entre}}{\sigma^2 + R\sigma_\epsilon^2} \sim \chi_{[(I-1)(J-1)]}^2$ independentemente de que $H_0^{(\Psi)}$ se verifique.

Portanto, adotando os mesmos argumentos do ítem a.3) verifica-se facilmente que a estatística

$$F = \frac{\frac{SQ(\hat{\Psi}_h)}{\sigma^2 + R\sigma_\varepsilon^2}}{1} = \frac{QM(\hat{\Psi}_h)}{QM_{Erro\ Entre}} \sim F_{[1, (I-1)(J-1)]} \quad (2.57)$$

e será a estatística de teste para testar a hipótese $H_0^{(\Psi)} : \Psi_h = 0$ vs $H_1^{(\Psi)} : \Psi_h \neq 0$. A regra de decisão será, rejeita $H_0^{(\Psi)}$, ao nível de significância α , se $F = \frac{QM(\hat{\Psi}_h)}{QM_{Erro\ Entre}} > F_{[1, (I-1)(J-1), \alpha]}$, em que $F_{[1, (I-1)(J-1), \alpha]}$ é o $100(1 - \alpha)$ -ésimo percentil superior da distribuição F com 1 grau de liberdade do numerador e $(I - 1)(J - 1)$ graus de liberdade do denominador.

- a.7) Para comparar os possíveis pares de médias da variável resposta relativas aos tratamentos, poderá ser empregado qualquer método de comparações múltiplas, embora neste trabalho deu-se preferência ao teste de Tukey, cuja teoria pode ser vista em (LEAL; PORRAS, 1998), (MONTGOMERY, 2007).

2.9.2 Hipóteses sobre Bloco

Tal como nas hipóteses sobre tratamento, estas também podem ser representada de duas maneiras, a saber:

$$a) \left\{ \begin{array}{l} H_0^{(\beta)} : \beta_j = 0, \forall j = 1, 2, \dots, J \\ vs \\ H_1^{(\beta)} : \beta_j \neq 0, \text{ para pelo menos um } \beta_j \end{array} \right. \quad (2.58)$$

ou

$$b) \left\{ \begin{array}{l} H_0^{(\beta)} : \mu_1 = \mu_2 = \dots = \mu_J = \mu \\ vs \\ H_1^{(\beta)} : \mu_j \neq \mu_{j'}, \text{ para pelo menos um par } (\mu_j, \mu_{j'}) \text{ } j \neq j' = 1, 2, \dots, J \end{array} \right.$$

Levando-se em conta os resultados obtidos no desenvolvimento da teoria e por argumentos semelhantes aos utilizados na seção 2.9.1, para as hipóteses sobre tratamento, pode-se verificar facilmente que

b.1) A estatística

$$F = \frac{\frac{SQ_{Bloco}}{\sigma^2 + R\sigma_\epsilon^2}}{(J-1)} = \frac{QM_{Bloco}}{QM_{Erro\ Entre}} \sim F_{[(J-1), (I-1)(J-1)]} \quad (2.59)$$

será a estatística de teste para testar a hipótese $H_0^{(\beta)} : \beta_j = 0$, para todo j , vs $H_0^{(\beta)} : \beta_j \neq 0$ para pelo menos um β_j . Rejeita-se $H_0^{(\beta)}$, ao nível de significância α , se $F = \frac{QM_{Bloco}}{QM_{Erro\ Entre}} > F_{[(J-1), (I-1)(J-1), \alpha]}$, em que $F_{[(J-1), (I-1)(J-1), \alpha]}$ é o $100(1 - \alpha)$ -ésimo percentil superior da distribuição F com $(J - 1)$ graus de liberdade do numerador e $(I - 1)(J - 1)$ graus de liberdade do denominador.

b.2) A estatística

$$F = \frac{\frac{SQ(\hat{\Psi}_h)}{\sigma^2 + R\sigma_\epsilon^2}}{1} = \frac{QM(\hat{\Psi}_h)}{QM_{Erro\ Entre}} \sim F_{[1, (I-1)(J-1)]} \quad (2.60)$$

será a estatística de teste para testar a hipótese de nulidade de um contraste h sobre as médias dos bocos, $H_0^{(\Psi)} : \Psi_h = 0$ vs $H_1^{(\Psi)} : \Psi_h \neq 0$. A regra de decisão será, rejeita $H_0^{(\Psi)}$, ao nível de significância α , se $F = \frac{QM(\hat{\Psi}_h)}{QM_{Erro\ Entre}} > F_{[1, (I-1)(J-1), \alpha]}$, em que $F_{[1, (I-1)(J-1), \alpha]}$ é o $100(1 - \alpha)$ -ésimo percentil superior da distribuição F com 1 grau de liberdade do numerador e $(I - 1)(J - 1)$ graus de liberdade do denominador.

b.3) As comparações múltiplas das médias dos blocos pelo teste Tukey, poderão ser feitas por meio do mesmo procedimento adotado no ítem a.7) com as seguintes modificações:

$$DMS = q_{[J; (I-1)(J-1); \alpha]} \sqrt{\frac{QM_{Erro\ Entre}}{IR}} \quad (2.61)$$

na qual $q_{[J; (I-1)(J-1); \alpha]}$ é o valor crítico da aptitude estudentizada de Tukey para $J =$ número de blocos envolvidos no ensaio, $(I - 1)(J - 1) =$ número de graus de liberdade do *Erro Entre* e nível de significância α (em geral $\alpha = 0,05$).

Em seguida calcular os valores absolutos das diferenças entre as estimativas dos possíveis pares de médias dos blocos envolvidos no experimento, $|\bar{y}_{.j} - \bar{y}_{.j'}|$, $j \neq j'$, $j, j' = 1, 2, \dots, J$. Finalmente, adotar a regra de decisão: Rejeiar H_0 em favor de H_1 , ao nível de significância α se, e somente se $|\bar{y}_{.j} - \bar{y}_{.j'}| > DMS$.

2.9.3 A tabela da ANOVA

A partir do conhecimento das partes componentes da variabilidade total representadas pelas somas de quadrados e das distribuições de probabilidade das estatísticas de teste das hipóteses sobre tratamento e Bloco deduzidas nas Seções 2.9.1 e 2.9.2, e, de acordo com Leal e Porras (1998) e Montgomery (2007), pode-se organizar a tabela completa da análise de variância - ANOVA, tal como se apresenta na literatura especializada, isto é,

Tabela 6: Tabela da Análise de Variância

F. Variação	GL	SQ	QM	F
Tratamento	$I - 1$	SQ_{Trat}	$QM_{Trat} = \frac{SQ_{Trat}}{I-1}$	$\frac{QM_{Trat}}{QM_{Erro\ Entre}}$
Bloco	$J - 1$	SQ_{Bloco}	$QM_{Bloco} = \frac{SQ_{Bloco}}{J-1}$	$\frac{QM_{Bloco}}{QM_{Erro\ Entre}}$
Erro Entre	$(I - 1)(J - 1)$	$SQ_{Erro\ Entre}$	$QM_{Erro\ Entre} = \frac{SQ_{Erro\ Entre}}{(I-1)(J-1)}$	-
Parcela	$IJ - 1$	$SQ_{Parcela}$	-	-
Erro Dentro	$IJ(R - 1)$	$SQ_{Erro\ Dentro}$	$QM_{Erro\ Dentro} = \frac{SQ_{Erro\ Dentro}}{IJ(R-1)}$	-
Total	$IJR - 1$	SQ_{Total}	-	-

3 Aplicação da teoria a um exemplo real

O objetivo deste Capítulo é apresentar uma aplicação da teoria desenvolvida neste trabalho a um conjunto de dados real recolhido de um experimento em blocos completos casualizados com repetições dos tratamentos dentro dos blocos.

3.1 Descrição do conjunto de dados experimentais

Para ilustrar o método, levou-se em conta um experimento analisado por Ferreira (1996), página 254, no qual considerou-se *três* variedades de cana-de-açúcar, *três* blocos e *três* repetições. A variável resposta analisada foi a porcentagem de açúcar provável, cujos valores observados encontram-se na Tabela 7, a seguir.

Tabela 7: Porcentagem de açúcar provável em variedades de cana-de-açúcar

Variedade	Repetição	Bloco			Soma
		1	2	3	
1	1	13,03	13,20	13,30	
	2	13,72	13,84	12,33	
	3	14,16	13,11	13,79	
		40,91	40,15	39,42	120,48
2	1	15,73	15,13	15,40	
	2	15,62	15,52	15,57	
	3	15,55	16,27	15,77	
		46,90	46,92	46,74	140,56
3	1	14,69	14,75	14,95	
	2	15,65	15,54	15,72	
	3	14,52	14,13	14,51	
		44,86	44,42	45,18	134,46
Soma		132,67	131,49	131,34	395,50

Fonte: Cortesia de Ms. Paulo Vanderlei Ferreira.

3.2 Cálculos das somas de quadrados e análise da variância

A partir dos dados da Tabela 7, calculou-se:

$$\begin{aligned} SQ_{Total} &= \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R y_{ijr}^2 - \frac{(y_{...})^2}{IJR} \\ &= (13,03)^2 + (13,72)^2 + \dots + (14,51)^2 - \frac{(395,5)^2}{3 \times 3 \times 3} = 29,3418; \end{aligned}$$

$$\begin{aligned} SQ_{Variedade} &= \frac{1}{JR} \sum_{i=1}^I y_{i..}^2 - \frac{(y_{...})^2}{IJR} \\ &= \frac{1}{3 \times 3} [(120,48)^2 + (140,56)^2 + (134,46)^2] - \frac{(395,5)^2}{3 \times 3 \times 3} = 23,5503; \end{aligned}$$

$$\begin{aligned} SQ_{Bloco} &= \frac{1}{IR} \sum_{j=1}^J y_{.j.}^2 - \frac{(y_{...})^2}{IJR} \\ &= \frac{1}{3 \times 3} [(132,67)^2 + (131,49)^2 + (131,34)^2] - \frac{(395,5)^2}{3 \times 3 \times 3} = 0,1179; \end{aligned}$$

$$\begin{aligned} SQ_{Parcela} &= \frac{1}{R} \sum_{i=1}^I \sum_{j=1}^J y_{ij.}^2 - \frac{(y_{...})^2}{IJR} \\ &= \frac{1}{3} [(40,91)^2 + (40,15)^2 + \dots + (45,18)^2] - \frac{(395,5)^2}{3 \times 3 \times 3} = 24,0239; \end{aligned}$$

$$\begin{aligned} SQ_{Erro\ Entre} &= SQ_{Parcela} - SQ_{Variedade} - SQ_{Bloco} \\ &= 24,0239 - 23,5503 - 0,1179 = 0,3557 \end{aligned}$$

e

$$\begin{aligned} SQ_{Erro\ Dentro} &= SQ_{Total} - SQ_{Parcela} \\ &= 29,3418 - 24,0239 = 5,3179. \end{aligned}$$

De posse dos resultados acima organiza-se a *Tabela da Análise de Variância - ANOVA*, de acordo como foi sugerido na Tabela 8, ou seja,

Tabela 8: Análise de variância para os dados da porcentagem de açúcar provável em variedades de cana-de-açúcar.

F. Variação	GL	SQ	QM	F
Variedade	2	23,5503	11,7752	132,45
Bloco	2	0,1179	0,0590	0,66
Erro Entre	4	0,3557	0,0889	-
Parcela	8	24,0239	3,0030	-
Erro Dentro	18	5,3179	0,2954	-
Total	26	29,3479	-	-

Como pode ser observado na Tabela 8, $F = 132,45$ e da tabela da distribuição F observa-se que $F_{[2;4;0,01]} = 18,00$. Como $F > F_{[2;4;0,01]}$, então rejeita-se a hipótese de igualdade das médias da porcentagem de açúcar provável relativas as variedades de cana-de-açúcar e concluí-se ao nível de significância $\alpha = 0,01$ que existe pelo menos um par de médias $(\mu_i, \mu_{i'})$ que diferem estatisticamente entre si.

Para verificar quais médias diferem entre si, adotou-se o seguinte procedimento:

1. Desdobrar os dois graus de liberdade de variedade em dois contrastes de interesse, cada um com um grau de liberdade. Imagine que são de interesse do pesquisador os seguintes contrastes representados pelas hipóteses $H_0^{(2)}$ e $H_0^{(3)}$ a seguir:

$$\left\{ \begin{array}{l} H_0^{(2)} : \mu_2 - \frac{\mu_1 + \mu_3}{2} = 0 \\ vs \\ H_1^{(2)} : \mu_2 - \frac{\mu_1 + \mu_3}{2} \neq 0 \end{array} \right. \text{ com 1 gl e } \left\{ \begin{array}{l} H_0^{(3)} : \mu_3 - \mu_1 = 0 \\ vs \\ H_1^{(3)} : \mu_3 - \mu_1 \neq 0 \end{array} \right. \text{ com 1 gl}$$

As estimativas desses contrastes são representadas por:

$$\hat{\Psi}_1 = \bar{y}_{2..} - \frac{\bar{y}_{1..} + \bar{y}_3}{2} = 15,62 - \frac{13,39 + 14,94}{2} = 1,46\%$$

e

$$\hat{\Psi}_2 = \bar{y}_{3..} - \bar{y}_{1..} = 14,94 - 13,39 = 1,55\%$$

Para facilitar os cálculos das respectivas somas de quadrados dos contrastes é conveniente reescrever as hipóteses $H_0^{(2)}$ e $H_0^{(3)}$ na forma equivalente, da seguinte maneira, $H_0^{(2)} : -\mu_1 + 2\mu_2 - \mu_3 = 0$ e $H_0^{(3)} : \mu_3 - \mu_1 = 0$. Daí, obtém-se:

$$SQ(\hat{\Psi}_1) = \frac{\left(\sum_i c_{1i} y_{i..} \right)^2}{JRK_1} = \frac{[(-1)(120,48)^2 + (2)(140,56) + (-1)(134,46)]^2}{3 \times 3 \times \underbrace{[(-1)^2 + (2)^2 + (-1)^2]}_{K_1}}$$

$$= \frac{(26, 18)^2}{54} = 12,6925 \text{ com 1 gl}$$

e

$$\begin{aligned} SQ(\hat{\Psi}_2) &= \frac{\left(\sum_i c_{2i} y_{i..}\right)^2}{JRK_2} = \frac{[(-1)(120, 48)^2 + (1)(134, 46)]^2}{3 \times 3 \times \underbrace{[(-1)^2 + (1)^2]}_{K_2}} \\ &= \frac{(13, 98)^2}{18} = 10,8578 \text{ com um gl.} \end{aligned}$$

Com estes resultados reorganiza-se a tabela da análise da variância, obtendo-se:

Tabela 9: Análise de variância para os dados da porcentagem de açúcar provável em variedades de cana-de-açúcar.

F. Variação	GL	SQ	QM	F
$H_0^{(2)} : \Psi_1 = \mu_2 - \frac{\mu_1 + \mu_3}{2} = 0$	1	12,6925	12,6925	142,77
$H_0^{(3)} : \Psi_2 = \mu_3 - \mu_1 = 0$	1	10,8578	10,8578	122,13
Variedade ($H_0^{(1)} = \mu_1 = \mu_2 = \mu_3 = \mu$)	(2)	(23,5503)	11,7752	132,45
Bloco	2	0,1179	0,0590	0,66
Erro Entre	4	0,3557	0,0889	-
Parcela	8	24,0239	3,0030	-
Erro Dentro	18	5,3179	0,2954	-
Total	26	29,3479	-	-

Com base nos resultados da Tabela 9 conclui-se que a média da porcentagem de açúcar provável da variedade V2 difere estatisticamente da média combinada das variedades V1 e V3, ao nível $\alpha = 0,01$ de significância, tendo em vista que $F_{[1; 4; 0,01]} = 21,20$. De modo análogo, pode-se concluir, ao nível $\alpha = 0,01$ de significância, que a porcentagem média de açúcar provável da variedade V3 difere estatisticamente da variedade V1.

2. Comparações das médias duas a duas pelo teste de Tukey:

Inicialmente calculou-se as estimativas das médias todas com desvio-padrão $s(\bar{y}_i)$ e a diferença mínima significativa -DMS, pelo método de Tukey ao nível $\alpha = 0,05$:

$$\begin{aligned} \bar{y}_1 &= 13,30\% \\ \bar{y}_2 &= 15,62\% \text{ todas com erro-padrão } s(\bar{y}_i) = \sqrt{\frac{QM_{\text{Erro Entre}}}{JR}} = 0,0994\%. \\ \bar{y}_3 &= 14,94\% \end{aligned}$$

e

$$\begin{aligned}
 DMS &= q_{[I; (I-1)(J-1); \alpha]} \sqrt{\frac{QM_{Erro\ Entre}}{JR}} \\
 &= q_{[3; 4; 0,05]} \sqrt{\frac{0,0889}{3 \times 3}} = 5,04 \sqrt{\frac{0,0889}{3 \times 3}} = 0,50\%;
 \end{aligned}$$

em seguida cauculou-se os valores absolutos das possíveis diferenças entre as estimativas das médias das variedades. Isto é,

$$\begin{aligned}
 |\bar{y}_{1..} - \bar{y}_{2..}| &= |13,39 - 15,62| = 2,23\% \\
 |\bar{y}_{1..} - \bar{y}_{3..}| &= |13,39 - 14,94| = 1,55\% \\
 |\bar{y}_{2..} - \bar{y}_{3..}| &= |15,62 - 14,94| = 0,68\%;
 \end{aligned}$$

Adotando-se a regra de decisão: rejeitar $H_0 : \mu_i = \mu_{i'}$, ao nível de significância $\alpha = 0,05$, se $|\bar{y}_{i..} - \bar{y}_{i'..}| > DMS$, conclui-se que as variedades de cana-de-açúcar estudadas produzem porcentagens médias de açúcar provável diferentes entre si.

3.3 Comprovação da idoneidade do modelo

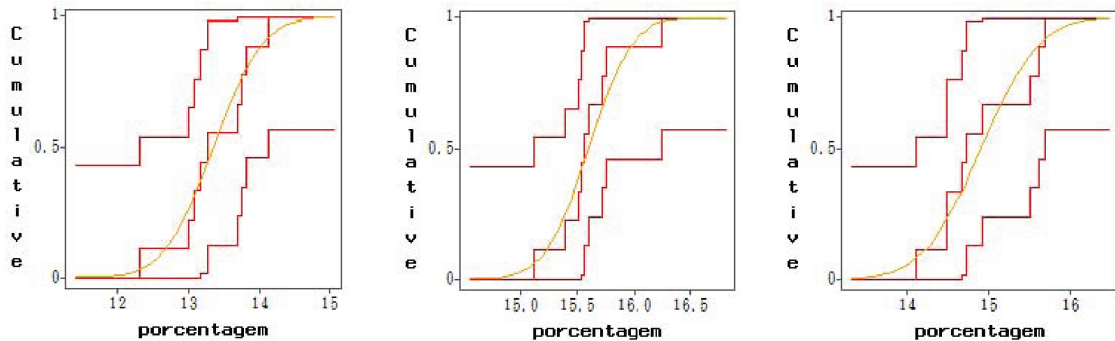


Figura 1: Função de distribuição acumulada sob normalidade por meio da estatística de Kolmogorov-Smirnov.

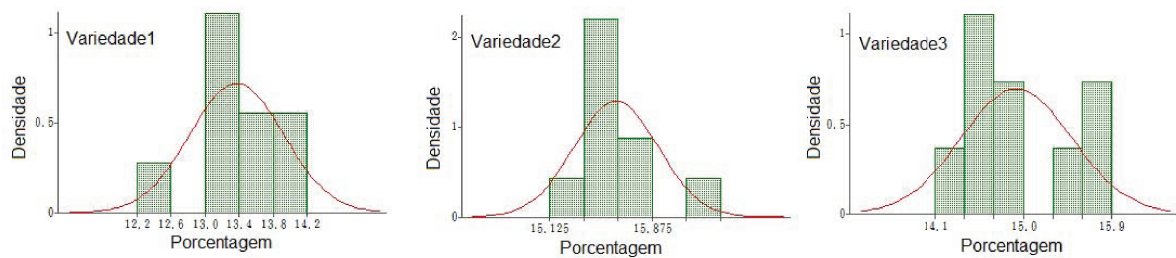


Figura 2: Histograma e polígono das frequências para cada variedade de cana-de-açúcar estudada.

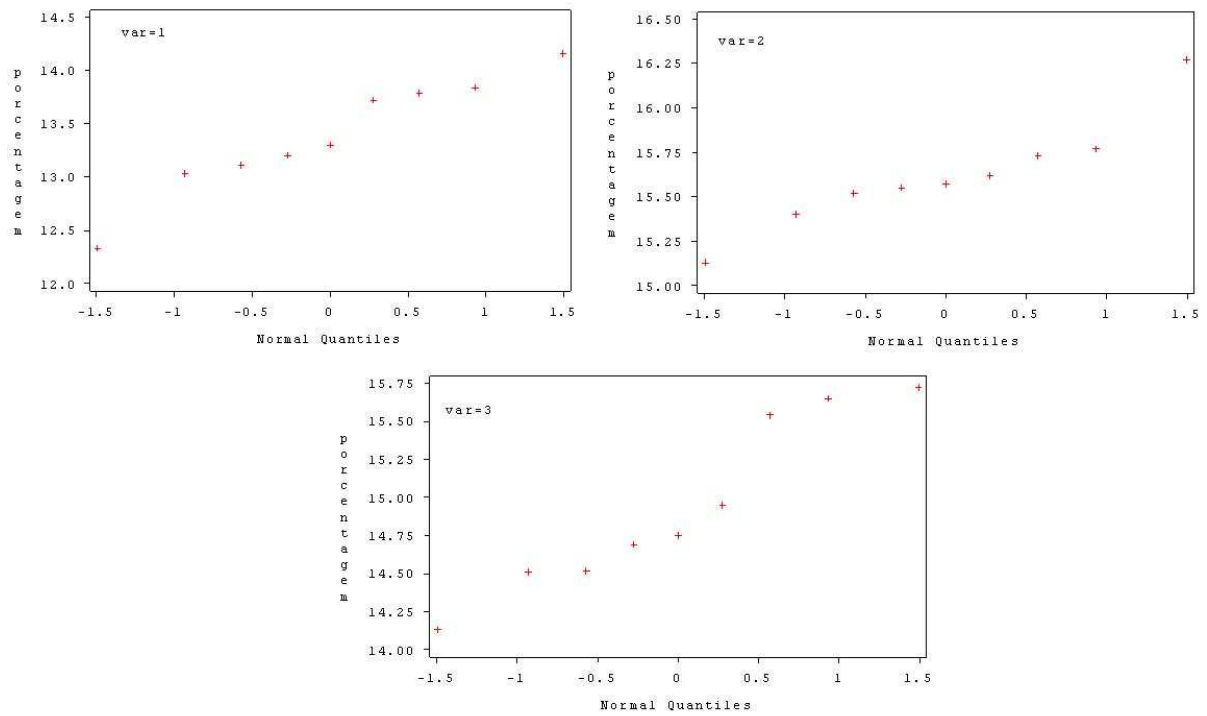


Figura 3: Valores plotados no gráfico Quantil para as variedades de cana-de-açúcar estudadas.

Conforme pode ser observado nas Figuras 1, 2 e 3 os erros são normalmente distribuídos.

A execução das análises pode ser facilitada empregando o *software* estatístico SAS por meio do seguinte procedimento:

```

OPTIONS NODATE PS=500;
DATA EXEMPLO;
DO VAR=1 TO 3;
  DO REP=1 TO 3;
    DO BLOCO=1 TO 3;
      INPUT PORCENTAGEM @@;
      OUTPUT;
    END;
  END;
END;
DATALINES;
13.03 13.20 13.30
13.72 13.84 12.33
14.16 13.11 13.79
15.73 15.13 15.40
15.62 15.52 15.57
15.55 16.27 15.77

```

```
14.69 14.75 14.95
15.65 15.54 15.72
14.52 14.13 14.51
;
RUN;
PROC PRINT DATA=EXEMPLO;
RUN;
PROC GLM DATA=EXEMPLO;
CLASS VAR BLOCO;
MODEL PORCENTAGEM=VAR BLOCO VAR*BLOCO / SS3;
TEST H=VAR BLOCO E=VAR*BLOCO;
CONTRAST "V2 vs (V1+V3)" VAR -1 2 -1;
CONTRAST "V3 vs V1" VAR -1 0 1;
MEANS VAR / TUKEY;
RUN;
```

4 Conclusão Final

As análises estatísticas para os dados de um experimento instalado num delineamento de *Blocos completos casualizados com repetições dos tratamentos dentro dos blocos*, consideram que as observações são representadas por um modelo matemático aditivo envolvendo uma média geral, os efeitos dos tratamentos, dos blocos e dois tipos de erros experimentais aleatórios: um, entre as unidades experimentais que receberam o mesmo tratamento em blocos diferentes, o *Erro Entre*, e outro, entre as unidades experimentais que receberam o mesmo tratamento dentro do mesmo bloco, o *Erro Dentro*. Neste delineamento, o *Erro Entre* pode se visto como o efeito da interação entre Tratamento e Bloco ($T \times B$). Além disso, supõe-se *a priori* que esses erros são independentemente distribuídos (um do outro e entre si) como uma normal de médias *zero* e variâncias comuns σ_ϵ^2 e σ^2 , respectivamente. Isto é, $\epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\epsilon^2)$, $\epsilon_{ijr} \stackrel{iid}{\sim} N(0, \sigma^2)$ e $Cov(\epsilon_{ij}, \epsilon_{ijr}) = 0$. Diante do exposto e após o desenvolvimento e aplicação da teoria que dar sustentação a estas análises, concluiu-se que:

1. Do ponto de vista prático, o planejamento de um experimento em blocos completos casualizados com repetição do conjunto de tratamentos nos blocos tem pouca aplicabilidade, uma vez que é apropriado a pesquisas envolvendo poucos tratamentos. E, do ponto de vista teórico, as contribuições tem sido muito poucas, dificultando de certa forma, a expansão das técnicas estatísticas envolvendo outras distribuições de probabilidade para os erros neste delineamento.
2. O método utilizado para estimar os parâmetros do modelo aqui adotado foi o dos *mínimos quadrados*. Preferiu-se esse método em detrimento do método da *Máxima verossimilhança* porque ele é mais simples e, sob normalidade, os dois métodos fornecem os mesmos estimadores, os quais têm excelentes propriedades;
3. Algumas estatísticas de teste apresentadas na Tabela 9 podem ser utilizadas para se fazer inferências marginais sobre, por exemplo, τ_i , μ_i , β_j e μ_j , embora, na prática, não sejam rotineiramente empregadas como ferramentas em busca de achados nas

pesquisas cientificamente planejadas nesse tipo de delineamento;

4. Não foi encontrado na literatura nenhum trabalho que justificasse completamente a base estatística que suporta as análises dos dados de um experimento em blocos completos casualizados com repetições dentro dos blocos. E, esta é, na opinião dos autores, a maior contribuição que se pode extrair deste trabalho.
5. É importante lembrar que em qualquer análise estatística, onde as observações foram obtidas a partir de um experimento cientificamente planejado, deve ser adotado como regra, a 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 da *Aditividade, Normalidade, Homocedasticidade e Independência dos erros*;
6. As análises estatísticas dos dados do experimento utilizado para ilustrar a teoria desenvolvida, apresentaram-se adequadas e constataram que a hipótese $H_0^{(\beta)} : \beta_j = 0$ não foi rejeitada, ao nível de significância $\alpha = 0,05$. Isto indica que os blocos não têm efeito sobre a porcentagem de açúcar provável nas variedades de cana-de-açúcar estudadas. Por outro lado, observou-se que a hipótese $H_0^{(\tau)} : \tau_i = 0, \forall i$, foi rejeitada ao nível de significância $\alpha = 0,01$. Daí, conclui-se que as variedades têm efeito sobre a porcentagem de açúcar provável (ou ainda, que existe pelo menos duas variedades que diferem entre si quanto a porcentagem média de açúcar provável). Ao testar a hipótese sobre o contraste entre a média de açúcar provável da variedade 2 contra a média combinada das variedades 1 e 3, verificou-se que este apresentou significância estatística. Finalmente, ao confrontar as porcentagens médias de açúcar provável das variedades pelo teste de Tukey, ao nível $\alpha = 0,05$, constatou-se que todas diferem estatisticamente entre si.

Referências

- BARBIN, D. *Componentes de Variância - Teoria e Aplicações*. 2. ed. Piracicaba - SP: FEALQ - Fundação de Estudos Agrários Luiz de Queiroz, 1993. 117 p.
- FERREIRA, P. V. *Estatística Experimental Aplicada à Agronomia*. Maceió - AL: Edufal, 1996.
- LEAL, J. G.; PORRAS, A. M. L. *Diseño estadístico de experimentos - Análises de la varianza*. Granada - ES: Grupo Editorial Universitario, 1998. 357 p.
- LOVE, H. H. *Introduction to Experimental Statistics*. New York: Mcgraw-Hill, 1964. 174 p.
- MONTGOMERY, D. C. *Design and Analysis of Experiments*. Singapore: Posts & Telecom Press, 2007. 642 p.
- ROHATGI, V. K. *An Introduction to Probability Theory and Mathematical Statistics*. New York: John Wiley, 1976. 684 p.
- ROUSSAS, G. G. *A Course in Mathematical Statistics*. 2. ed. New York: Academic Press, 1997. 572 p.
- SILVA, J. G. C. da. Estatística experimental: Planejamento de experimentos. Versão Preliminar, p. 130–154, 2007.