**COEFICIENTE DE CORRELAÇÃO
INTRACLASSE:**

Planejamento com Alocação Ótima e
Aplicação no Estudo de
Confiabilidade de Medidas

Ting Hui Ching

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE MESTRE
EM
ESTATÍSTICA

Área de Concentração: Estatística
Orientador: **Prof. Dr. Clóvis de Araújo Peres**

Este trabalho foi financiado pela CNPQ

- São Paulo, Julho de 1995 -

**COEFICIENTE DE CORRELAÇÃO
INTRACLASSE:**

Planejamento com Alocação Ótima e
Aplicação no Estudo de
Confiabilidade de Medidas

Este exemplar corresponde à redação
final da dissertação devidamente
corrigida e apresentada por
Ting Hui Ching
e aprovada pela Comissão Julgadora.

São Paulo, 23 de outubro de 1995.

Banca Examinadora

. Prof. Dr. Clóvis de Araújo Peres (Orientador)
. Profa. Dra. Carmen Diva Saldiva de André
. Prof. Dr. José Ruy Porto de Carvalho

- IME-USP
- IME-USP
- NIIA/EMBRAPA

Agradecimentos

Foi uma longa caminhada, mas finalmente conseguimos!

Durante esse tempo, aprendi bastante, não somente os conhecimentos científicos, mas também as lições de vida. Senti fortemente a presença de Deus em todos os momentos; sem ele, não teria este trabalho de hoje. Graças a ele, tive privilégios para me dedicar a este trabalho, sem me preocupar muito com a saúde, com a família e tudo mais. E também graças a ele, contei com a ajuda de muitas pessoas, entre as quais agradeço especialmente:

Ao meu orientador, Prof. Dr. Clóvis de Araújo Peres, pela orientação, paciência, compreensão e incentivo constante.

À grande amiga Lurdes Y. T. Inoue pela amizade, pela revisão do texto, pelas sugestões e pelo apoio nos momentos mais difíceis.

À Rosângela H. Loschi (minha amiga Bina) pelos conselhos, incentivo, e pelas valiosas discussões.

À Profa. Lisbeth K. Cordani pela atenção e apoio.

À Dra. Elizabeth In Myung Kim do Hospital Universitário da USP pelas discussões valiosas na aplicação prática.

Ao casal Chiann e Anil pela amizade e apoio.

À minha família pela compreensão e apoio constante.

Aos amigos por sempre acreditarem em mim e pela paciência.

E às pessoas (conhecidas e desconhecidas) que calorosamente me acolheram nos momentos em que precisei.

Abstract

The Intraclass Correlation Coefficient (ρ) is an important parameter associated to one-way random effects model. It is widely applied in different fields of research. In Genetic, for example, ρ is used for estimating the heritability of selected traits in animal and plant populations. In Psychology, it plays an important role in reliability theory. In Epidemiology, it is commonly used to measure the degree of familial resemblance with respect to biological or environmental characteristics. And in Sensitivity Analysis, ρ is used to measure the effectiveness of an experimental treatment.

The variance of estimators of ρ strongly depends on the selected design, with respect to the number of factor classes and the number of replications within each class. In this work, we study designs which minimize the variance of estimators of the intraclass correlation coefficient and other parameters of interest. The application of ρ in the study of reliability of measurement is discussed. We show some consequences of unreliability as functions of ρ and present some solutions.

Sumário

O Coeficiente de Correlação Intraclassa (ρ) é uma medida importante ligada ao modelo de efeitos aleatórios com um fator. A sua aplicação se encontra em amplas áreas de pesquisa. Em Genética, por exemplo, ρ é utilizado para estimar a herdabilidade de alguma característica na população animal ou vegetal; em Psicologia, é uma medida importante no estudo de confiabilidade; em Epidemiologia, mede o grau de semelhança familiar por fator biológico ou característica ambiental; em Análise de Sensibilidade, ρ é usado para medir a eficiência de um tratamento experimental.

A variância dos estimadores de ρ dependem fortemente do planejamento utilizado, no que se refere ao número de classes do fator e ao número de réplicas em cada classe. Neste trabalho, apresentamos planejamentos com alocação ótima que tornam mínimas as variâncias dos estimadores do coeficiente de correlação intraclassa e de outros parâmetros de interesse (também definidos através do modelo de efeitos aleatórios com um fator). Apresentamos também a aplicação de ρ no estudo de confiabilidade de medidas. Expressamos algumas consequências da não-confiabilidade como funções de ρ , e apresentamos algumas soluções.

Aos meus pais e
ao meu irmão

Conteúdo

Introdução	1
1 Modelo de Efeitos Aleatórios com Um Fator: Alguns Resultados	5
1.1 Desenvolvimento Histórico	7
1.2 Modelo	10
1.3 Notação Matricial	11
1.4 Partições da Soma de Quadrados	15
1.5 Distribuições das Somas de Quadrados	16
1.5.1 Distribuição de SQA	16
1.5.2 Distribuição de SQR	17
1.5.3 Independência das Distribuições	18
1.6 Estimação Pontual dos Componentes de Variância e Suas Funções	19
1.7 Construção de Intervalos de Confiança	21
1.7.1 Intervalo de Confiança para σ_a^2	21
1.7.2 Intervalo de Confiança para θ	21
1.7.3 Intervalo de Confiança para ρ	22
1.8 Testes de Hipóteses	23
2 Planejamento Ótimo na estimação de σ_a^2 e $\theta = \frac{\sigma_a^2}{\sigma_e^2}$	24
2.1 Estimação de σ_a^2	25
2.1.1 Variância de $\widehat{\sigma_a^2}$ expressa em função de θ	26

2.1.2	Variância de $\widehat{\sigma}_a^2$ expressa em função de ρ	30
2.2	Estimação de θ	34
2.2.1	Variância de $\widehat{\theta}$ expressa em função de θ	35
2.2.2	Variância de $\widehat{\theta}$ expressa em função de ρ	38
2.3	Uso do programa PO.COM	40
3	Planejamento Ótimo na Estimação de $\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$	43
3.1	Estimador de Análise de Variância	44
3.2	Estimador de Máxima Verossimilhança	49
4	Coeficiente de Correlação Intraclasse como Medida de Confiabilidade	54
4.1	Confiabilidade e ρ	55
4.2	Estimação da confiabilidade de uma medida	57
4.3	Algumas Conseqüências da Não-Confiabilidade	63
4.3.1	Sub-estimação da Correlação	63
4.3.2	Problemas na Análise Estatística de Ensaios Clínicos	64
4.4	Controle da Confiabilidade pelo Número de Réplicas	69
4.5	Incorporando o custo	70
4.6	Estudo de Confiabilidade Inter-juízes	73
4.6.1	Modelo com Dois Fatores Aleatórios	75
4.6.2	Modelo Misto - com um fator fixo	77
4.7	Aplicação	79
5	Considerações Finais	84
	Apêndice I	86
	Apêndice II	88
	Apêndice III	90
	Bibliografia	99

Introdução

"O processo de preparação para enfrentar o futuro, visando assumir os riscos certos e aproveitar as oportunidades que ele oferece, tornando-os mais favoráveis, chama-se Planejamento." - Antônio Maximiano

Diz o provérbio: "Partindo-se de um bom começo, já temos meio caminho andado." Ora, planejamento é uma forma de bom começo, pois tanto na nossa vida cotidiana, quanto no campo da pesquisa científica, um bom planejamento sempre nos traz a vantagem de economizar recursos, tais como tempo, dinheiro e mão-de-obra. Em outras palavras, ele permite que o objetivo seja alcançado com maior eficiência. Nosso intuito, neste trabalho, é o de fornecer aos colegas da nossa área de estudo algumas ferramentas úteis, que permitam a otimização do recurso (o que em muitas vezes é limitado), com o ganho de eficiência e precisão nos resultados, quando o modelo de efeitos aleatórios é adotado.

O modelo básico para o nosso trabalho é o modelo de efeitos aleatórios com um fator. Este modelo é utilizado quando a variável resposta do estudo é uma variável contínua, e as unidades experimentais são sorteadas por um esquema de dupla aleatorização. Apresentamos este esquema na Figura 0.1.

Nessa figura, cada coluna representa uma classe (nível), que é sorteada na população das classes do fator em estudo, e temos n_i réplicas em cada classe (quando os números das réplicas dentro das classes forem iguais, dizemos que o modelo tem dados balanceados), que também são sorteadas entre todas as possíveis réplicas de cada classe. Em aplicações da área biológica, freqüentemente, as classes são representadas por indivíduos ou famílias; na área industrial, as

Figura 0.1: Estrutura de dados para o modelo de efeitos aleatórios com um fator

Classe	1	2	..	k
	x	x	..	x
	x	x	..	x
	x	x	..	x
	x	x	..	x

classes são, em geral, representadas por lotes, bateladas ("batches"), etc.

Neste modelo, a variação das observações pode ser decomposta em duas partes, a variação devido a classe e a variação devido às réplicas, denotadas por σ_a^2 e σ_e^2 , respectivamente. Em geral, quando este modelo é adotado, o interesse reside na estimação dos dois componentes de variância ou algumas funções deles.

Dentro deste contexto, o alvo principal desse estudo é o coeficiente de correlação intraclass (ρ), definido por $\frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$. Este parâmetro, desde a sua introdução por Fisher(1925), tem recebido a atenção de pesquisadores de várias áreas distintas.

Nesse trabalho, vamos mostrar como determinar um PLANEJAMENTO ÓTIMO que permita estimar o coeficiente de correlação intraclass com máxima precisão, quando o número total de unidades experimentais (fator que influencia a variância do estimador) é fixado de antemão por consideração de custo. Apresentamos também a aplicação de ρ no estudo de confiabilidade de medidas (com o intuito de efetuar um controle sobre a qualidade dos dados de uma pesquisa). Adicionalmente, apresentamos o planejamento ótimo para estimar 2 outros parâmetros importantes - a variância das classes (σ_a^2) e a razão entre os componentes de variância ($\theta = \frac{\sigma_a^2}{\sigma_e^2}$) - definidos pelo mesmo modelo.

O método que adotamos para determinar o planejamento ótimo na estimação dos parâmetros é o mesmo utilizado por Scheffé(1959) na estimação de σ_a^2 e θ . Este método é estendido a ρ . Visando facilitar a interpretação de alguns resultados, incorporamos alguns gráficos. Além

disso, para facilitar a busca do planejamento ótimo, elaboramos um programa computacional baseado nos procedimentos estudados.

O conteúdo do trabalho é apresentado em 5 capítulos e 3 apêndices. No capítulo 1, apresentamos o nosso modelo básico - o modelo de efeitos aleatórios com 1 fator . Iniciamos este capítulo com 2 exemplos para ilustrar as diferenças¹ básicas entre este modelo e o modelo de efeitos fixos. Alguns resultados de estimação pontual (pelo método de Análise de Variância) e estimação por intervalo dos 3 parâmetros de interesse são apresentados para fornecer alguns elementos básicos para a leitura dos outros capítulos. Apresentamos também, neste capítulo, um resumo acerca do desenvolvimento histórico do modelo e a simplificação dos cálculos na obtenção das distribuições das somas de quadrados utilizando o Produto de Kronecker².

O segundo e terceiro capítulos enfocam a busca do planejamento ótimo na estimação dos parâmetros de interesse. No capítulo 2, estudamos o planejamento ótimo na estimação de σ_a^2 e θ . Apresentamos também a utilização do programa computacional PO.COM³ por nós elaborado, que facilita a busca do planejamento ótimo. No capítulo 3, elaboramos o planejamento ótimo na estimação de ρ . Como ρ é o parâmetro sobre o qual reside o nosso principal interesse, além da busca do planejamento ótimo utilizando o estimador de Análise de Variância, estudamos também o planejamento ótimo com outro estimador usual de ρ - o estimador de Máxima Verossimilhança.

No capítulo 4, apresentamos a aplicação de ρ no estudo de confiabilidade de medidas. Iniciamos este capítulo com a definição do conceito de confiabilidade e o relacionamos com ρ . As estimações, pontual e por intervalo, de ρ são apresentadas com uma interpretação dentro de um novo contexto, isto é, o de confiabilidade. Fatos importantes como as conseqüências da não-confiabilidade são explorados neste capítulo de maneira que nos possibilite perceber ainda mais a importância de ρ . Apresentamos também um exemplo real aplicando os resultados vistos no capítulo.

¹Um resumo das diferenças, incluindo definição e problemas inferenciais, pode ser encontrado no Apêndice I.

²Definição e propriedades do Produto de Kronecker podem ser encontradas no Apêndice II.

³A listagem e as especificações do programa podem ser encontradas no Apêndice III.

Finalmente, no último capítulo , fazemos alguns comentários gerais deste trabalho e sugerimos uma importante extensão dos resultados.

Capítulo 1

Modelo de Efeitos Aleatórios com Um Fator: Alguns Resultados

A ação de transcrever situações práticas (fenômenos físicos) em equações matemáticas é chamada *construção de modelos*. Quando um modelo é linear nos parâmetros, este modelo pertence à família de modelos lineares.

O *Modelo de Efeitos Aleatórios* e o *Modelo de Efeitos Fixos* são membros dessa família. A equação matemática que descreve os fenômenos explicados pelos dois modelos são iguais, a diferença existe, no entanto, nas suposições associadas a cada modelo.

A escolha entre os dois modelos depende do objetivo do estudo. Quando o interesse é quantificar os efeitos e fazer a comparação entre eles, o modelo de efeitos fixos é utilizado. Quando o interesse é o de conhecer a variação entre os efeitos e não os seus valores individuais, o modelo de efeitos aleatórios deve ser adotado.

O interesse principal do nosso estudo está associado com o modelo de efeitos aleatórios, porém achamos importante mostrar ao leitor algumas idéias sobre o uso dos dois modelos, para que a sua escolha seja a mais conveniente. Vamos ilustrar essas idéias através de 2 exemplos simples. Consideremos a seguinte situação:

Exemplo 1: Em uma indústria fabricante de máquinas que produzem parafusos, deseja-se veri-

ficar a qualidade das máquinas fabricadas. Uma maneira de avaliar se as máquinas estão em boa condição de funcionamento é, verificar se a resistência dos parafusos produzidos está de acordo com a especificação pré-determinada pela indústria. k máquinas são selecionadas aleatoriamente para este teste, e a resistência é verificada através de uma amostra de n parafusos por máquina.

Podemos descrever este fenômeno com a seguinte equação:

$$Y_{ij} = \mu + A_i + \epsilon_{ij}, \quad i = 1, 2, \dots, k \quad j = 1, 2, \dots, n$$

onde

Y_{ij} é a resistência do parafuso j produzido pela máquina i ,

μ é a resistência média dos parafusos,

A_i é o efeito da máquina i ,

ϵ_{ij} é o erro aleatório da observação Y_{ij} .

Esperamos que não exista a variação da resistência dos parafusos devido às máquinas, isto é, que a variância de A_i seja zero. Como A_i provém de uma população de máquinas, suponhamos que a distribuição dos efeitos A_i tem média zero e variância σ_a^2 , e o efeito de uma máquina é independente do efeito de uma outra máquina, ou seja, os efeitos das máquinas são aleatórios e independentes entre si. Os erros aleatórios da equação também são supostos ter uma distribuição com média zero, variância σ_e^2 , e são independentes entre si e dos efeitos da máquina.

Com este conjunto de informações (a equação e as suposições), definimos o modelo de efeitos aleatórios com um fator (neste exemplo, o fator é *máquina*) em que o interesse está na variância dos efeitos.

Consideremos, agora, outra situação:

Exemplo 2: Em uma indústria de parafusos, existem k máquinas na linha de produção. O objetivo do controle rotineiro é o de verificar se a resistência dos parafusos produzidos está de acordo com a sua especificação. O processo consiste em selecionar uma amostra de cada

máquina e avaliar a resistência de cada parafuso. A máquina que não produz conforme o padrão sofrerá uma regulagem.

Esta situação pode ser descrita pela mesma equação do modelo de efeitos aleatórios. A diferença é de que, as máquinas neste exemplo são as únicas da indústria, e estamos interessados em verificar o efeito de cada uma. Portanto, a suposição sobre os A_i 's no exemplo anterior não é mais válida neste caso. Por questão de estimação, a restrição $\sum A_i = 0$ é adicionada a equação. A suposição sobre os erros aleatórios continua a mesma.

Com essas informações, define-se o modelo de efeitos fixos com um fator.

A descrição formal das diferenças entre os dois modelos pode ser encontrada no Apêndice I.

O modelo básico do nosso estudo é o modelo de efeitos aleatórios. Este modelo também chamado de modelo de componentes de variância, é amplamente utilizado nas mais diversas áreas de pesquisa.

Neste capítulo, vamos apresentar alguns resultados importantes para o modelo com um fator aleatório (com dados balanceados), utilizando a notação do Produto de Kronecker (Apêndice II). Iniciamos o capítulo com o desenvolvimento histórico do modelo de efeitos aleatórios. Na seção 1.2, apresentamos o modelo com um fator aleatório e os parâmetros de interesse do nosso trabalho. Na seção 1.3, apresentamos este modelo na forma matricial. Na seção 1.4, definimos as somas de quadrados e as escrevemos em forma quadrática utilizando o Produto de Kronecker. Este procedimento facilitará na obtenção das distribuições das somas de quadrados, que é apresentada na seção 1.5. Nas seções 1.6 e 1.7, apresentamos a estimação pontual (usando o método de Análise de Variância) e intervalos de confiança para os parâmetros de nosso interesse. Na última seção, apresentamos os testes de hipóteses usuais deste modelo.

1.1 Desenvolvimento Histórico

Modelos de efeitos aleatórios aparecem na literatura desde o século XIX. Os astrônomos da época, Airy (1861) e Chauvenet (1863) usaram o modelo de efeitos aleatórios com um fator

em seus trabalhos¹. Porém, foi Fisher (1918) quem introduziu o termo "análise de variância" na literatura, iniciando uma nova era para este modelo.

Em 1925, o livro de Fisher, *Statistical Methods for Research Workers*, trouxe outro avanço na teoria dos componentes de variância. Neste livro, Fisher introduziu o *Coefficiente de Correlação Intraclass* no modelo com um fator, e utilizou o método de análise de variância(ANOVA) para estimá-lo. Ele estendeu os resultados do modelo balanceado para o modelo não-balanceado e o modelo de 2 fatores cruzados com interação. Os conceitos básicos de vários métodos de estimação são também apresentados neste livro.

Fisher(1925) formalizou o método de análise de variância, porém foi Tippett (1931), que relacionou o modelo linear com esse método, e o usou para explicar os conceitos da análise de variância. Ele também foi o primeiro a estender este método de estimação para um modelo de 2 fatores sem considerar interação.

Sobre o uso dos componentes de variância, Yates & Zaccopani (1935) fizeram uma aplicação pioneira em um experimento com cereais.

Comparado com o modelo de dados balanceados, o modelo com dados não-balanceados é bem mais complexo. Isso dificultou o desenvolvimento deste assunto. De fato, após a introdução de Fisher, levou quase 15 anos para surgirem os trabalhos de Cochran(1939) e Winsor & Clarke (1940) que discutem a estimação dos componentes de variância através da ANOVA para modelos com um fator com dados não-balanceados.

A estimação dos componentes de variância pelo método de Máxima Verossimilhança sofreu a mesma negligência que os dados não-balanceados e pela mesma razão - complexidade. O método foi introduzido também por Fisher em 1925, mas só após quatro décadas, com o trabalho de Hartley & Rao (1967), o uso deste método foi estimulado. No artigo, eles propuseram um procedimento iterativo em busca da solução para equações de verossimilhança, do modelo generalizado misto com dados não-balanceados (*general unbalanced mixed model*).

Sob a suposição de normalidade, Crump (1947) calculou a variância amostral dos esti-

¹Uma interpretação moderna do modelo de Airy(1861) pode ser encontrada em Scheffé(1956).

madores dos componentes de variância para modelos com um e com dois fatores cruzados. A variância amostral dos estimadores sem a suposição de normalidade, foi calculada mais tarde por Hammersley(1949). Neste trabalho, ele fez também um estudo de planejamento com alocação ótima que aumenta a precisão na estimação dos componentes de variância isolados.

Crump(1947) e Ganguli(1941) notaram uma deficiência dos estimadores da ANOVA, isto é, esses podem produzir estimativas negativas, o que segundo a definição, não podia acontecer, e propuseram a solução de substituir a estimativa negativa por zero².

Entre os primeiros trabalhos que associam modelos lineares com os componentes de variância (Neyman et al. 1935, Daniels 1939, Jackson 1939), Eisenhart (1947) foi o primeiro a definir precisamente as diferenças entre os modelos com efeitos fixos e aleatórios, e introduziu a terminologia Modelo I e Modelo II para os dois modelos, respectivamente.

Até 1952 os resultados sobre componentes de variância na literatura eram dispersos. Anderson & Bancroft(1952) foi o primeiro livro a formalizar e apresentar os resultados de forma didática.

A literatura ligada ao modelo de efeitos aleatórios é imensa. Khuri & Sahai (1985) fizeram uma revisão bibliográfica selecionada sobre estimação e testes de hipóteses neste modelo e, ela fornece uma visão bem completa sobre este assunto. Uma lista abrangente de aproximadamente 2.700 citações sobre tópicos relacionados com modelos de efeitos aleatórios publicados até 1983 é apresentada em Sahai, Khuri & Kapadia (1985) e Sahai(1979). Informações básicas sobre estimação e teste de hipóteses em modelos de efeitos aleatórios são também encontrados em livros-textos como Brownlee (1965) e Jonhson & Leone(1964) com aplicação na área industrial, Anderson & Bancroft(1952), Scheffé (1959), Graybill (1961), Searle (1971), Rao (1973), Box et al(1978), Neter et al(1990) e outros.

²Para maiores informações a este respeito, recomendamos a dissertação de mestrado de Bezerra(1976).

1.2 Modelo

O modelo aleatório com um fator para experimentos completamente casualizados e balanceados, é usualmente expresso por

$$Y_{ij} = \mu + A_i + \epsilon_{ij} \quad i = 1, \dots, k; \quad j = 1, \dots, n \quad (1.1)$$

onde:

Y_{ij} : j-ésima observação da i-ésima classe;

μ : média geral da população;

A_i : efeito aleatório da i-ésima classe;

ϵ_{ij} : erro aleatório da j-ésima observação da i-ésima classe;

e $kn = N$ é o número total de observações.

As suposições usuais do modelo são:

i) $A_i \sim Normal(0, \sigma_a^2)$;

ii) $\epsilon_{ij} \sim Normal(0, \sigma_e^2)$;

iii) A_i e ϵ_{ij} são completamente independentes.

Com estas suposições, a esperança de Y_{ij} é

$$E(Y_{ij}) = E(\mu + A_i + \epsilon_{ij}) = \mu$$

a variância de Y_{ij} é

$$Var(Y_{ij}) = Var(\mu + A_i + \epsilon_{ij}) = \sigma_a^2 + \sigma_e^2 = \sigma_y^2, \quad (1.2)$$

isto é, a soma de 2 componentes, de onde decorre a nomenclatura modelo de componentes de variância.

Ao contrário do que acontece com o modelo de efeitos fixos, neste modelo, as observações não são completamente independentes. As observações apresentam a seguinte estrutura de covariância:

Para observações j e j' da mesma classe i

$$Cov(Y_{ij}, Y_{ij'}) = E(Y_{ij} - \mu)(Y_{ij'} - \mu) = E(A_i + \epsilon_{ij})(A_i + \epsilon_{ij'}) = E(A_i^2) = \sigma_a^2 \quad (1.3)$$

e, para as observações de classes diferentes,

$$Cov(Y_{ij}, Y_{i'j}) = E(A_i + \epsilon_{ij})(A_{i'} + \epsilon_{i'j}) = 0$$

Quando este modelo é adotado, o interesse principal é o de estimar os componentes σ_a^2 e σ_e^2 - que denominamos de variância das classes e variância dos erros aleatórios, respectivamente - ou suas funções. Uma função dos componentes de variância que é muito utilizada em áreas biológicas ou médicas é o coeficiente de correlação intraclasse (ρ), o qual é definido por

$$\rho = \frac{\sigma_a^2}{\sigma_y^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}. \quad (1.4)$$

Pela (1.2) e (1.3) segue que,

$$\rho = \frac{Cov(Y_{ij}, Y_{ij'})}{Var(Y_{ij})},$$

ou seja, ρ também é a correlação usual das observações de uma mesma classe.

A razão entre os componentes de variância (θ) é outra função importante dos componentes de variância, que é definida por

$$\theta = \frac{\sigma_a^2}{\sigma_e^2}, \quad (1.5)$$

parâmetro que é de grande interesse na genética e na indústria química, por exemplo.

1.3 Notação Matricial

Com a notação matricial, o modelo apresentado em (1.1) toma a forma:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}^*$$

onde $\mathbf{Y}$ é o vetor das observações; $\mathbf{X}$ a matriz de planejamento; $\boldsymbol{\beta}$ vetor dos parâmetros e $\boldsymbol{\epsilon}^*$ é o vetor dos efeitos aleatórios; com

$$Y_{N \times 1} = \begin{pmatrix} y_{11} \\ y_{12} \\ \cdot \\ \cdot \\ y_{1n} \\ y_{21} \\ y_{22} \\ \cdot \\ \cdot \\ y_{2n} \\ \cdot \\ \cdot \\ y_{k1} \\ y_{k2} \\ \cdot \\ \cdot \\ y_{kn} \end{pmatrix}$$

$$\mathbf{X}_{N \times k} = \begin{pmatrix} 1 & 0 & . & . & 0 \\ 1 & 0 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & 0 & . & . & 0 \\ 0 & 1 & 0 & . & 0 \\ 0 & 1 & 0 & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 1 & 0 & 0 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & 1 \\ 0 & 0 & . & . & 1 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & 1 \end{pmatrix}$$

$$\boldsymbol{\beta}_{k \times 1} = \begin{pmatrix} \mu \\ \mu \\ . \\ . \\ \mu \end{pmatrix}$$

$$\epsilon_{N \times 1}^* = \mathbf{A} + \epsilon = \begin{pmatrix} A_1 \\ \cdot \\ \cdot \\ A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_2 \\ \cdot \\ \cdot \\ A_k \\ \cdot \\ \cdot \\ A_k \end{pmatrix} + \begin{pmatrix} \epsilon_{11} \\ \cdot \\ \cdot \\ \epsilon_{1n} \\ \epsilon_{21} \\ \cdot \\ \cdot \\ \epsilon_{2n} \\ \cdot \\ \cdot \\ \epsilon_{k1} \\ \cdot \\ \cdot \\ \epsilon_{kn} \end{pmatrix}$$

Com as suposições de normalidade e independência apresentadas na seção anterior,

$$\mathbf{Y} \sim \text{Normal}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

onde $\boldsymbol{\mu}$ é o vetor de médias de dimensão $N \times 1$; $\boldsymbol{\Sigma}$ é a matriz de covariância

$$\boldsymbol{\Sigma}_{N \times N} = \begin{pmatrix} V & 0 & 0 & \dots & 0 \\ 0 & V & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & V \end{pmatrix}$$

com

$$V_{n \times n} = \begin{pmatrix} \sigma_a^2 + \sigma_e^2 & \sigma_a^2 & \sigma_a^2 & \dots & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 + \sigma_e^2 & \sigma_a^2 & \dots & \sigma_a^2 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2 & \dots & \sigma_a^2 + \sigma_e^2 \end{pmatrix}$$

Utilizando o Produto de Kronecker podemos escrever

$$\Sigma = I_k \otimes V$$

onde I_k é a matriz identidade de ordem k .

1.4 Partições da Soma de Quadrados

A soma de quadrados total pode ser decomposta em duas parcelas correspondentes às fontes de variações do modelo.

$$\sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - \bar{Y})^2 = \sum_{i=1}^k n(\bar{Y}_{i.} - \bar{Y})^2 + \sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - \bar{Y}_{i.})^2$$

$$SQT = SQA + SQR$$

SQA se refere à variação devido ao fator e SQR representa a variação residual. Estas somas de quadrados podem ser expressas como formas quadráticas em termos do Produto de Kronecker, o que torna muito simples a determinação das distribuições de probabilidade de cada uma delas.

$$SQA = Y'(X(X'X)^{-1}X' - \frac{1}{N}E_N)Y = Y'(\frac{1}{n}I_k \otimes E_n - \frac{1}{N}E_N)Y = Y'M_A Y$$

$$SQR = Y'(I_N - X(X'X)^{-1}X')Y = Y'(I_N - \frac{1}{n}I_k \otimes E_n)Y = Y'M_R Y$$

$$SQT = Y'(I_N - \frac{1}{N}E_N)Y$$

onde: E_N é uma matriz quadrada, formada somente por elementos iguais a 1, I_N é a matriz identidade, e X é a matriz de planejamento definida anteriormente (os sub-índices representam a dimensão das matrizes). Então temos

$$Y'(I_N - \frac{1}{N}E_N)Y = Y'(\frac{1}{n}I_k \otimes E_n - \frac{1}{N}E_N)Y + Y'(I_N - \frac{1}{n}I_k \otimes E_n)Y$$

1.5 Distribuições das Somas de Quadrados

Quando os dados são balanceados, as somas de quadrados podem ser facilmente escritas em termos do Produto de Kronecker. Desta forma, torna-se simples a obtenção das distribuições das somas de quadrados³.

Para obter as distribuições das somas de quadrados utilizamos o seguinte teorema:

Teorema 1 (Rao, 1973) *Se X tem distribuição Normal(μ , Σ), então $X'AX$ tem distribuição Qui-Quadrado não-central (χ_r^2), se e somente se Σ for positiva definida e $A\Sigma A = A$; onde $r = \text{Posto}(A)$ é o número de graus de liberdade, e $\lambda = \frac{1}{2}\mu' A \mu$ é o parâmetro de não-centralidade.*

1.5.1 Distribuição de SQA

Na seção 1.4, vimos que

$$SQA = Y' M_A Y$$

onde,

$$M_A = [\frac{1}{n}I_k \otimes E_n - \frac{1}{N}E_N] = \frac{1}{n}[I_k - \frac{1}{k}E_k] \otimes E_n$$

Observamos que:

$$M_A \Sigma M_A = [\frac{1}{n}(I_k - \frac{1}{k}E_k) \otimes E_n][I_k \otimes V][\frac{1}{n}(I_k - \frac{1}{k}E_k) \otimes E_n]$$

³No caso de dados não-balanceados, não conseguimos escrever as somas de quadrados com o Produto de Kronecker, e as distribuições delas geralmente não são conhecidas.

$$M_A \Sigma M_A = \left[\frac{1}{n^2} (I_k - \frac{1}{k} E_k) I_k (I_k - \frac{1}{k} E_k) \right] \otimes E_n V E_n$$

Como $E_n V E_n = n(\sigma_e^2 + n\sigma_a^2) E_n$ e $I_k - \frac{1}{k} E_k$ é idempotente, temos:

$$M_A \Sigma M_A = (\sigma_e^2 + n\sigma_a^2) \frac{1}{n} (I_k - \frac{1}{k} E_k) \otimes E_n = (\sigma_e^2 + n\sigma_a^2) M_A$$

Temos também que,

$$\lambda = \mu' M_A \mu = 0$$

e,

$$\text{posto}(M_A) = \text{Tr}(M_A) = \text{Tr}\left[\frac{1}{n} (I_k - \frac{1}{k} E_k) \otimes E_n\right] = \text{Tr}\left[\frac{1}{n} (I_k - \frac{1}{k} E_k)\right] \times \text{Tr}(E_n) = k - 1$$

Decorre do Teorema 1 que

$$SQA \sim (\sigma_e^2 + n\sigma_a^2) \chi_{k-1}^2(\text{central}) \quad (1.6)$$

1.5.2 Distribuição de SQR

Temos que:

$$SQR = Y' M_R Y$$

onde,

$$M_R = (I_N - \frac{1}{n} I_k \otimes E_n) = (I_k \otimes I_n - \frac{1}{n} I_k \otimes E_n) = [I_k \otimes (I_n - \frac{1}{n} E_n)]$$

Observamos que:

$$M_R \Sigma M_R = [I_k \otimes (I_n - \frac{1}{n} E_n)] [I_k \otimes V] [I_k \otimes (I_n - \frac{1}{n} E_n)]$$

$$M_R \Sigma M_R = I_k \otimes (I_n - \frac{1}{n} E_n) V (I_n - \frac{1}{n} E_n)$$

Como $E_n V = V E_n$ e $E_n V E_n = n V E_n$,

$$M_R \Sigma M_R = I_k \otimes (I_n - \frac{1}{n} E_n) V = I_k \otimes \sigma_e^2 (I_n - \frac{1}{n} E_n)$$

$$\mathbf{M}_R \boldsymbol{\Sigma} \mathbf{M}_R = \sigma_e^2 [\mathbf{I}_k \otimes (\mathbf{I}_n - \frac{1}{n} \mathbf{E}_n)] = \sigma_e^2 \mathbf{M}_R$$

Além disso,

$$\lambda = \boldsymbol{\mu}' \mathbf{M}_R \boldsymbol{\mu} = 0$$

e,

$$\text{posto}(\mathbf{M}_R) = \text{posto}[\mathbf{I}_k \otimes (\mathbf{I}_n - \frac{1}{n} \mathbf{E}_n)] = \text{Tr}(\mathbf{I}_k) \text{Tr}(\mathbf{I}_n - \frac{1}{n} \mathbf{E}_n) = k \times (n - 1) = N - k$$

Portanto,

$$SQR \sim \sigma_e^2 \chi_{N-k}^2(\text{central}) \quad (1.7)$$

1.5.3 Independência das Distribuições

A independência estatística entre SQA e SQR pode ser verificada através do seguinte teorema:

Teorema 2 (Searle, 1971) Quando

$$\mathbf{X} \sim \text{Normal}(\boldsymbol{\mu}, \mathbf{V}),$$

as formas quadráticas $\mathbf{X}'\mathbf{A}\mathbf{X}$ e $\mathbf{X}'\mathbf{B}\mathbf{X}$ tem distribuições independentes se e somente se $\mathbf{A}\mathbf{V}\mathbf{B}=0$ (ou $\mathbf{B}\mathbf{V}\mathbf{A}=0$).

Em geral, este teorema é aplicado em situações onde as formas quadráticas têm distribuições χ^2 , porém, isto não é um pré-requisito para seu uso. Não é necessário nos preocuparmos com as distribuições das formas quadráticas, basta que $\mathbf{X}$ seja uma variável com distribuição normal.

Assim, para verificar a independência entre SQA e SQR, precisamos mostrar que $\mathbf{M}_A \boldsymbol{\Sigma} \mathbf{M}_R = 0$ ou $\mathbf{M}_R \boldsymbol{\Sigma} \mathbf{M}_A = 0$.

Temos,

$$\begin{aligned} \mathbf{M}_R \boldsymbol{\Sigma} \mathbf{M}_A &= [\mathbf{I}_k \otimes (\mathbf{I}_n - \frac{1}{n} \mathbf{E}_n)] [\mathbf{I}_k \otimes \mathbf{V}] [\frac{1}{n} (\mathbf{I}_k - \frac{1}{k} \mathbf{E}_k) \otimes \mathbf{E}_n] \\ \mathbf{M}_R \boldsymbol{\Sigma} \mathbf{M}_A &= [\frac{1}{n} (\mathbf{I}_k - \frac{1}{k} \mathbf{E}_k) \otimes (\mathbf{I}_n - \frac{1}{n} \mathbf{E}_n) \mathbf{V} \mathbf{E}_n] = [\frac{1}{n} (\mathbf{I}_k - \frac{1}{k} \mathbf{E}_k) \otimes 0] = 0 \end{aligned}$$

portanto, SQA e SQR são estatisticamente independentes.

1.6 Estimação Pontual dos Componentes de Variância e Suas Funções

Vamos utilizar o método da Análise de Variância (ANOVA) para estimar os componentes de variância. Este método é bastante adotado por sua simplicidade, e produz, sob a suposição de normalidade, os estimadores não-viesados de variância uniformemente mínima (ENVVUM)(Graybill & Wortham,1956). Mesmo sem esta suposição, eles continuam ser ENVVUM na classe dos estimadores quadráticos (Graybill & Hultquist, 1961). Os estimadores são as soluções das equações formadas pela igualdade que é estabelecida entre os quadrados médios e as suas esperanças.

Teorema 3 (Searle,1971) *Seja χ_r^2 uma variável com distribuição de Qui-quadrado central, com r graus de liberdades, então*

$$E(\chi_r^2) = r, Var(\chi_r^2) = 2r.$$

Portanto,

$$E(QMA) = E\left(\frac{SQA}{k-1}\right) = E\left(\frac{\sigma_e^2 + n\sigma_a^2}{k-1}\chi_{k-1}^2\right) = \sigma_e^2 + n\sigma_a^2$$

$$E(QMR) = E\left(\frac{SQR}{N-k}\right) = E\left(\frac{\sigma_e^2}{N-k}\chi_{N-k}^2\right) = \sigma_e^2$$

Podemos resumir os resultados na seguinte tabela:

Tabela 1.1: Tabela de Análise de Variância para o Modelo de Efeitos Aleatórios com um fator e dados balanceados.

Fonte de Variação	Graus de Liberdade	Soma de Quadrados	Quadrado Médio	Esperança de Quadrados Médios
Entre Classes	k-1	SQA	$QMA = \frac{SQA}{k-1}$	$\sigma_e^2 + n\sigma_a^2$
Dentre classes	N-k	SQR	$QMR = \frac{SQR}{N-k}$	σ_e^2

Os estimadores de σ_a^2 e σ_e^2 obtidos pelo método da Análise de Variância são dados por:

$$\widehat{\sigma_e^2} = QMR \quad (1.8)$$

$$\widehat{\sigma_a^2} = \frac{QMA - QMR}{n} \quad (1.9)$$

Substituindo (1.8) e (1.9) nas definições de ρ e θ apresentadas em (1.4) e (1.5), podemos obter os estimadores de θ e ρ .

$$\widehat{\theta} = \frac{QMA - QMR}{nQMR} \quad (1.10)$$

$$\widehat{\rho} = \frac{QMA - QMR}{QMA + (n - 1)QMR} \quad (1.11)$$

Como θ e ρ não são funções lineares dos componentes de variâncias, os estimadores assim produzidos são, em geral, viesados.

Decorrente de suas definições, os componentes de variância e suas funções devem assumir somente valores positivos. No entanto, os estimadores apresentados em (1.9), (1.10) e (1.11) podem fornecer estimativas negativas. Isto ocorre devido a uma propriedade intrínseca do método e não a uma inadequação do modelo ou invalidação das suposições. Uma sugestão comum para a estimativa negativa é, a sua substituição pelo valor zero. Outras soluções alternativas podem ser encontradas em Searle(1971, pag 407-408).

Além dos estimadores da ANOVA, existem outros tipos de estimadores tais como, os de máxima verossimilhança (MV), os de Bayes, etc. Vários estudos de comparação de alguns estimadores envolvendo modelos diferentes foram feitos (Khuri & Sahai,1985 e Donner, 1986). Entre os quais, o estudo de simulação de Swallow & Monahan(1984) motrou que, o estimador de ANOVA de σ_a^2 tem um bom comportamento, comparando seu erro quadrático médio e viés com os de outros estimadores. Donner & Koval(1980) comparou vários estimadores de ρ e concluiu que, para os valores pequenos ($\rho \leq 0,1$) e grandes ($\rho \geq 0,8$), a acurácia do estimador de máxima verossimilhança é maior, enquanto para os valores intermediários, os estimadores de ANOVA e de MV têm quase a mesma acurácia. Em geral, o estimador de ANOVA é uma boa opção.

1.7 Construção de Intervalos de Confiança

O intervalo de confiança nos fornece uma margem de incerteza e nos permite julgar qual a possível magnitude do erro que estamos cometendo. Nesta seção vamos apresentar a construção dos intervalos de confiança dos parâmetros σ_a^2 , θ e ρ no caso de dados balanceados⁴.

1.7.1 Intervalo de Confiança para σ_a^2

A distribuição de $\hat{\sigma}_a^2$ é determinada através da diferença de duas distribuições χ^2 que nem sempre é χ^2 (Searle, 1971). Há vários métodos para a construção de intervalo de confiança para σ_a^2 propostos na literatura e, seja no caso de dados balanceados ou não, a maioria deles são aproximados (Khuri & Sahai, 1985). Apresentamos aqui o resultado de Bulmer (1957).

O método desenvolvido por Bulmer (1957) produz o seguinte intervalo, com coeficiente de confiança $1 - \alpha$.

$$L_{Inf} = \frac{QMR}{n} \left[\frac{(k-1)QMA}{\chi_{k-1, 1-\alpha/2}^2 QMR} - 1 + \frac{F_{k-1, N-k, 1-\alpha/2}}{\frac{QMA}{QMR}} \left(1 - \frac{(k-1)F_{k-1, N-k, 1-\alpha/2}}{\chi_{k-1, 1-\alpha/2}^2} \right) \right]$$

$$L_{Sup} = \frac{QMR}{n} \left[\frac{(k-1)QMA}{\chi_{k-1, \alpha/2}^2 QMR} - 1 + \frac{F_{k-1, N-k, \alpha/2}}{\frac{QMA}{QMR}} \left(1 - \frac{(k-1)F_{k-1, N-k, \alpha/2}}{\chi_{k-1, \alpha/2}^2} \right) \right]$$

1.7.2 Intervalo de Confiança para θ

Quando os dados do modelo são balanceados, pela distribuição de $\hat{\theta}$, podemos construir um intervalo exato para ele.

Sabemos que

$$\mathcal{F} = \frac{QMA}{QMR} \sim \left(1 + n \frac{\sigma_a^2}{\sigma_e^2} \right) F_{k-1, N-k, p} = (1 + n\theta) F_{k-1, N-k, p},$$

onde $F_{k-1, N-k, p}$ é a distribuição de Fisher-Snedecor com $k-1$ e $N-k$ graus de liberdades, e probabilidade p na cauda direita.

⁴Quando os dados são balanceados, os intervalos constuídos para θ e ρ são exatos, caso contrário, são aproximados.

Podemos determinar o intervalo de confiança para θ , de modo que:

$$P[(1 + n \frac{\sigma_a^2}{\sigma_e^2})F_1 < \frac{QMA}{QMR} < (1 + n \frac{\sigma_a^2}{\sigma_e^2})F_2] = 1 - \alpha$$

$$P(\frac{QMA}{QMR} \frac{1}{F_2} < 1 + n \frac{\sigma_a^2}{\sigma_e^2} < \frac{QMA}{QMR} \frac{1}{F_1}) = 1 - \alpha$$

$$P[\frac{1}{n}(\frac{QMA}{QMR} \frac{1}{F_2} - 1) < \theta < \frac{1}{n}(\frac{QMA}{QMR} \frac{1}{F_1} - 1)] = 1 - \alpha$$

Então

$$L_i = \frac{1}{n}(\frac{QMA}{QMR} \frac{1}{F_2} - 1), \quad (1.12)$$

é o limite inferior para θ , com $F_2 = F_{k-1, N-k, 1-\alpha/2}$. E

$$L_s = \frac{1}{n}(\frac{QMA}{QMR} \frac{1}{F_1} - 1) \quad (1.13)$$

é o limite superior desse intervalo, com $F_1 = F_{k-1, N-k, \alpha/2}$

1.7.3 Intervalo de Confiança para ρ

Uma forma de se obter um intervalo de confiança para ρ é considerá-lo como função de θ , e a partir daí, gerar os limites, inferior e superior, do intervalo.

Dividindo o denominador e numerador em (1.4) por σ_e^2 , escrevemos ρ em função de θ :

$$\rho = \frac{\theta}{1 + \theta}$$

O intervalo de confiança para ρ é

$$P(\frac{L_i}{1 + L_i} \leq \rho \leq \frac{L_s}{1 + L_s}) = 1 - \alpha$$

onde L_i e L_s são as expressões apresentadas em (1.12) e (1.13), respectivamente.

1.8 Testes de Hipóteses

No modelo de efeitos aleatórios com um fator, duas hipóteses de interesse muito comum são: $H_0 : \sigma_a^2 = 0$ e $H_0 : \sigma_a^2 \leq \theta_0 \sigma_e^2$ (onde $\theta_0 \geq 0$, é uma constante conhecida). Para testar a primeira hipótese, usamos a estatística

$$\mathcal{F} = \frac{QMA}{QMR} = \frac{\frac{SQA}{k-1}}{\frac{SQR}{N-k}}$$

Como

$$SQR \sim \sigma_e^2 \chi_{N-k}^2,$$

$$SQA \sim (\sigma_e^2 + n\sigma_a^2) \chi_{k-1}^2,$$

e são independentes, então

$$\mathcal{F} \sim (1 + n \frac{\sigma_a^2}{\sigma_e^2}) F_{k-1, N-k, p},$$

onde F é a distribuição F de Fisher-Snedecor central.

A hipótese nula é rejeitada, se $\mathcal{F} > F_{k-1, N-k, p}$.

Para a segunda hipótese $H_0 : \sigma_a^2 \leq \theta_0 \sigma_e^2$, a estatística tem distribuição

$$\mathcal{F} = \frac{QMA}{QMR} \sim (1 + n\theta_0) F_{k-1, N-k, p}$$

então,

$$\frac{1}{1 + n\theta_0} \frac{QMA}{QMR} \sim F_{k-1, N-k, p}$$

e a hipótese nula é rejeitada, se $\frac{1}{1 + n\theta_0} \frac{QMA}{QMR} > F_{k-1, N-k, p}$.

Capítulo 2

Planejamento Ótimo na estimação de σ_a^2 e $\theta = \frac{\sigma_a^2}{\sigma_e^2}$

Vimos no Capítulo 1 que, quando o modelo de efeitos aleatórios é adotado, o interesse reside na estimação dos componentes de variância ou algumas funções dos componentes que têm interpretação dentro do problema em estudo. Para o modelo com um fator, os parâmetros de interesse mais comuns são μ , σ_a^2 , σ_e^2 , θ e ρ definidos no capítulo anterior.

Para experimentos balanceados, o fator do modelo tem k classes, com n unidades experimentais em cada classe do fator, totalizando $N (= kn)$ unidades experimentais. Em muitos problemas reais, por razão de custo, o número total de unidades experimentais é fixado antes da realização do experimento. Neste caso, para cada valor de N fixado podemos ter várias escolhas de n e k . Como a variância dos estimadores dos parâmetros de interesse é função de n e k , podemos encontrar valores de n e k que minimizam esta variância. Este critério de otimização do planejamento é referido na literatura como planejamento com alocação ótima ("optimum allocation design"), ou simplesmente, planejamento ótimo.

Adotando o modelo de efeitos aleatórios com um fator e dados balanceados, neste capítulo, vamos encontrar o planejamento com alocação ótima quando os parâmetros de interesse a serem estimados (utilizando o método de ANOVA) são σ_a^2 e θ .

Além da alocação das unidades experimentais, a variância dos estimadores depende de uma função dos componentes de variância, θ ou ρ . Dependendo da área em estudo, às vezes, as informações preliminares sobre θ são mais fáceis de serem obtidas do que sobre ρ , e vice-versa. Para podermos aplicar os resultados do planejamento ótimo em diferentes áreas, neste estudo, vamos apresentar as variâncias dos estimadores como funções de θ e de ρ .

Scheffé(1959) e Peres(1990) apresentaram o estudo de alocação ótima como função de θ . O mesmo procedimento será adotado aqui, e vamos estender o estudo para a alocação ótima como função de ρ .

O conteúdo deste capítulo é apresentado em 3 seções. Na seção 2.1, apresentamos o estudo de alocação ótima na estimação do parâmetro σ_a^2 e, na seção 2.2, de θ . Cada uma das duas primeiras seções é subdividida em 2 subseções. Na primeira subseção, os resultados são apresentados como funções de θ , e na segunda subseção como funções de ρ . Para facilitar a visualização, os resultados mais importantes são acompanhados pela representação gráfica. Na última seção, apresentamos a utilização do programa PO.COM, interpretando os resultados fornecidos por ele. Este programa foi elaborado, com base em toda a teoria estudada neste capítulo, para facilitar a busca do planejamento ótimo. A listagem do programa pode ser encontrada no Apêndice III.

2.1 Estimação de σ_a^2

Já vimos no capítulo anterior que o estimador da análise de variância para σ_a^2 , não-viesado, com variância uniformemente mínima, é dado pela seguinte expressão:

$$\widehat{\sigma_a^2} = \frac{1}{n}[QMA - QMR].$$

Mostramos na seção 1.5.3 que QMA e QMR são independentes. Decorrente desse fato, a variância do estimador de σ_a^2 é:

$$Var(\widehat{\sigma_a^2}) = Var\left[\frac{1}{n}(QMA - QMR)\right] = \frac{1}{n^2}[Var(QMA) + Var(QMR)],$$

como $QMA = \frac{SQA}{k-1} \sim \frac{\sigma_e^2 + n\sigma_a^2}{k-1} \chi_{k-1}^2$, e $QMR = \frac{SQR}{N-k} \sim \frac{\sigma_e^2}{N-k} \chi_{N-k}^2$, pelo Teorema 3 segue que

$$Var(\widehat{\sigma_a^2}) = \frac{1}{n^2} [2(k-1) \left(\frac{\sigma_e^2 + n\sigma_a^2}{k-1} \right)^2 + 2(N-k) \left(\frac{\sigma_e^2}{N-k} \right)^2].$$

2.1.1 Variância de $\widehat{\sigma_a^2}$ expressa em função de θ

Podemos escrever $Var(\widehat{\sigma_a^2})$ como função de N , k e $\theta = \frac{\sigma_a^2}{\sigma_e^2}$, de modo que

$$Var(\widehat{\sigma_a^2}) = 2\sigma_e^4 \left[\frac{\theta^2}{k-1} + \frac{2\theta k}{N(k-1)} + \frac{(N-1)k^2}{N^2(k-1)(N-k)} \right], \quad 1 < k < N. \quad (2.1)$$

Em estudos de alocação ótima, dizemos que a precisão de $\widehat{\sigma_a^2}$ é máxima quando a variância apresentada em (2.1) for mínima para os valores de N e σ_e^2 fixos e o parâmetro θ conhecido. Neste caso, a única variável em (2.1) é k , e o valor de k que conduz a uma variância mínima é denominado k_{otimo} .

Em busca do valor k_{otimo} , estudamos primeiro o comportamento da $Var(\widehat{\sigma_a^2})$ como função de k , considerando que k varia continuamente no intervalo $(1, N)$.

Como

$$\lim_{k \rightarrow 1} Var(\widehat{\sigma_a^2}) = \lim_{k \rightarrow N} Var(\widehat{\sigma_a^2}) = \infty$$

e os termos que compõem $Var(\widehat{\sigma_a^2})$ são contínuos e positivos, $Var(\widehat{\sigma_a^2})$ tem pelo menos um mínimo no intervalo $(1, N)$ (Veja a Figura 2.1).

Para determinar este(s) ponto(s) de mínimo, fazemos

$$\frac{dVar(\widehat{\sigma_a^2})}{dk} = 0,$$

e obtemos a equação

$$k^2(N^2\theta^2 + 2N\theta - N^2 + 1) - 2k(N^3\theta^2 + 2N^2\theta - N^2 + N) + (N^4\theta^2 + 2N^3\theta) = 0$$

As raízes desta equação são

$$k_1 = \frac{N^2\theta}{N\theta - N + 1} \quad e \quad (2.2)$$

$$k_2 = \frac{N^2\theta + 2N}{N\theta + N + 1}. \quad (2.3)$$

Figura 2.1: $Var(\widehat{\sigma}_a^2)$ como função de k para $N=40$, $\sigma_e^2=1$ e θ fixo

Vamos estudar as raízes k_1 e k_2 como funções de θ . A raiz cujo contradomínio tem intersecção com o intervalo $(1,N)$ é ponto de mínimo, isto é, é um ponto que minimiza a $Var(\widehat{\sigma}_a^2)$.

i) Estudo de k_1

Podemos escrever $k_1 = \frac{N\theta}{\theta - (1 - \frac{1}{N})}$. Verificamos que, se $\theta < 1 - \frac{1}{N}$, k_1 é negativo; se $\theta > 1 - \frac{1}{N}$, $k_1 > N$; e se $\theta = 1 - \frac{1}{N}$, k_1 é indefinido;. O contradomínio de k_1 não intercepta o intervalo $(1,N)$, portanto, não é o ponto que minimiza $Var(\widehat{\sigma}_a^2)$. A variação de k_1 em função de θ é apresentada na Figura 2.2.

ii) Estudo de k_2

Para N fixo, k_2 é uma função crescente em θ . Tende a $\frac{2N}{N+1}$ quando θ tende a zero, e tende a N quando θ tende a infinito (veja a Figura 2.3).

Como o contradomínio de k_2 está contido no intervalo $(1,N)$, esta é a raiz que minimiza $Var(\widehat{\sigma}_a^2)$, isto é, para qualquer valor de θ fixado em (2.3), o número de classes k_2 obtido é ótimo, no sentido em que, a sua adoção no experimento conduz a uma precisão máxima do estimador da ANOVA para os valores de N e θ fixados. Denotaremos k_2 por k_{otimo} .

Figura 2.2: Solução k_1 da primeira derivada de $\text{Var}(\widehat{\sigma}_a^2)$ como função de θ para $N=40$

Figura 2.3: k_{otimo} que minimiza $\text{Var}(\widehat{\sigma}_a^2)$ como função de θ com diversos valores de N

Vimos antes que o número de classes (k) pode assumir qualquer valor no intervalo $1 < k < N$. Porém, do ponto de vista prático, k só pode assumir valor inteiro no intervalo $2 \leq k \leq \frac{N}{2}$, pois se $k > \frac{N}{2}$, o único número inteiro que n pode assumir é o valor um, o que impossibilita a estimação de σ_e^2 . Além disso, o valor de k deverá ser ajustado para que $\frac{N}{k}$ (número de réplicas) seja inteiro.

Para que k_{otimo} obedeça a restrição prática de que $2 \leq k \leq \frac{N}{2}$, θ deve assumir valores apenas no seguinte intervalo,

$$\frac{2}{N(N-2)} \leq \theta \leq \frac{N-3}{N}.$$

Assim, nos casos em que $\theta < \frac{2}{N(N-2)}$, a melhor escolha¹ é $k = 2$ e, conseqüentemente, o número de observações em cada classe é $n = \frac{N}{2}$. Quando $\theta > \frac{N-3}{N}$, a melhor escolha² é $k = \frac{N}{2}$ e o número de observações em cada classe é $n = 2$. Para $\frac{2}{N(N-2)} \leq \theta \leq \frac{N-3}{N}$, a melhor escolha de k é k_{otimo} dado em (2.3). Como a obtenção de k_{otimo} depende de um parâmetro desconhecido θ , podemos substituir θ na expressão (2.3) por uma estimativa baseada na experiência ou intuição (Scheffé, 1959)³. Se o k_{otimo} assim obtido for divisor de N , adotamos este valor como solução e obtemos o planejamento ótimo tendo $k = k_{otimo}$ e $n = \frac{N}{k_{otimo}}$. Caso contrário, k é, entre os dois divisores de N mais próximos de k_{otimo} , aquele que produz a menor $Var(\hat{\sigma}_a^2)$. Geralmente, o divisor escolhido é aquele que tem a menor distância ao k_{otimo} ; porém, quando as distâncias dos dois divisores ao k_{otimo} são aproximadamente iguais, as variâncias têm que ser comparadas para a escolha do planejamento ótimo.

Exemplo: Se um experimento anterior nos informa que θ é aproximadamente igual a 0,25, e o pesquisador quer realizar o experimento com $N=60$ unidades experimentais, o pla-

¹Como não é possível utilizar k_{otimo} , o planejamento com $k=2$ é, entre todos os possíveis planejamentos, aquele que produz a variância mais próxima do planejamento com $k=k_{otimo}$.

²idem.

³Segundo Peres(1990), esta dependência de k_{otimo} em θ não traz nenhum problema quando σ_a^2 é maior do que σ_e^2 , pois θ é maior do que 1, e portanto $\theta > \frac{N-3}{N}$. Neste caso, a solução ótima é $k = \frac{N}{2}$, como observamos antes. Esta situação ocorre com frequência em aplicações práticas.

planejamento com k_{otimo} é aquele em que $k=13,42$ e $n = \frac{N}{k} = 4,47$. Como k não é divisor de N , devemos escolher um divisor de N mais próximo de k , ou seja, 12 ou 15. Como $Var(\widehat{\sigma}_a^2|k=12) = Var(\widehat{\sigma}_a^2|k=15)$, tanto $k=12$, quanto $k=15$ pode ser utilizado. Portanto, temos dois planejamentos ótimos: um em que $k=12$ e $n=5$ e, outro onde $k=15$ e $n=4$.

Resumidamente, se conhecemos o valor de θ ou temos alguma estimativa para ele, então podemos calcular k_{otimo} utilizando a expressão (2.3).

Se conhecemos somente a relação de ordem entre σ_a^2 e σ_e^2 , então também podemos determinar o planejamento ótimo. Para $\sigma_a^2 < \frac{2}{N(N-2)}\sigma_e^2$ (isto é, σ_a^2 pequeno em relação a σ_e^2), o planejamento ótimo é $k=2$ e $n = \frac{N}{2}$, enquanto para $\sigma_a^2 > \frac{(N-3)}{N}\sigma_e^2$ (isto é, σ_a^2 maior que σ_e^2), o planejamento ótimo é $k = \frac{N}{2}$ e $n=2$.

Contudo, nem sempre conhecemos θ , ou uma relação de ordem entre σ_a^2 e σ_e^2 , e nem sempre é possível conseguir uma estimativa confiável de θ . Neste caso, não podemos calcular k_{otimo} , e portanto não conseguimos determinar o planejamento ótimo. Visando contornar este problema, aproveitamos a idéia de adoção do divisor mais próximo do k_{otimo} (quando k_{otimo} não é divisor de N), na elaboração de um programa computacional PO.COM. Este programa apresenta todos os possíveis planejamentos de uma amostra de tamanho N , com os valores de θ correspondentes a cada mudança de planejamento. Então, mesmo desconhecendo o valor exato de θ , o pesquisador pode também escolher o seu planejamento ótimo, procurando o provável intervalo a que θ pertence no arquivo de saída fornecido pelo programa. Exemplos podem ser encontrados na seção 2.3.

2.1.2 Variância de $\widehat{\sigma}_a^2$ expressa em função de ρ

Em algumas áreas, é mais fácil obter informações preliminares sobre o parâmetro ρ (definido na seção 1.2) do que sobre o parâmetro θ . Para que os resultados de planejamento ótimo obtidos na subseção anterior possam também ser utilizados nessa área de pesquisa, nesta seção, vamos expressar os resultados encontrados na seção 2.1.1 como função de ρ .

Para expressar $Var(\widehat{\sigma}_a^2)$ como função de ρ , basta substituir na expressão (2.1): $\theta = \frac{\rho}{1-\rho}$. Obtemos:

$$Var(\widehat{\sigma}_a^2) = 2\sigma_e^4 \left[\frac{\rho^2}{(k-1)(1-\rho)^2} + \frac{2\rho k}{N(k-1)(1-\rho)} + \frac{(N-1)}{N^2} \frac{k^2}{(k-1)(N-k)} \right], \quad 1 < k < N \quad (2.4)$$

A função apresentada em (2.4) tem um comportamento semelhante àquela da expressão (2.1). Existe pelo menos um ponto que minimiza esta função. Para determiná-lo(s), repetimos o procedimento utilizado na seção 2.1.1 e obtemos como possíveis soluções ótimas

$$k_1 = \frac{N^2 \rho}{1 - N - \rho + 2N\rho} \quad e \quad (2.5)$$

$$k_2 = \frac{N(2 - 2\rho + N\rho)}{1 + N - \rho}. \quad (2.6)$$

Notemos que estas soluções podem ser obtidas também substituindo $\theta = \frac{\rho}{1-\rho}$ nas raízes das expressões (2.2) e (2.3) da seção 2.1.1.

Vamos estudar as raízes k_1 e k_2 como funções de ρ . A raiz que apresenta contradomínio $(1, N)$ é um ponto de mínimo.

i) Estudo de k_1

Podemos escrever $k_1 = \frac{N^2 \rho}{(1-N) - \rho(1-2N)}$. É fácil de se verificar que, se $\rho > \frac{1-N}{1-2N}$, $k_1 > N$; se $\rho < \frac{1-N}{1-2N}$, k_1 é negativo; e se $\rho = \frac{1-N}{1-2N}$, k_1 é indefinido (veja a Figura 2.4). Como o contradomínio de k_1 não contém o intervalo $(1, N)$, k_1 não é a raiz que minimiza $Var(\widehat{\sigma}_a^2)$.

ii) Estudo de k_2

Para N fixo, k_2 é uma função crescente em ρ (veja a Figura 2.5). Para $0 \leq \rho \leq 1$, o contradomínio de k_2 é $[\frac{2N}{1+N}, N]$, e está contido no intervalo $(1, N)$. Portanto, k_2 é a raiz que minimiza $Var(\widehat{\sigma}_a^2)$. Denotamos k_2 por k_{otimo} .

Para obtermos o valor de k_{otimo} que minimiza $Var(\widehat{\sigma}_a^2)$ de uma amostra de tamanho N , substituímos o valor de ρ em (2.6) por estimativas obtidas de experimentos anteriores, ou por intuição.

Figura 2.4: solução k_1 da primeira derivada de $Var(\widehat{\sigma}_a^2)$ como função de ρ para $N=40$

Figura 2.5: k_{otimo} que minimiza $Var(\widehat{\sigma}_a^2)$ como função de ρ com diversos valores de N

A adoção de k_{otimo} no experimento induz a uma menor variância do estimador; isto é,

$$Var(\widehat{\sigma}_a^2 | k_{otimo}) = 2\sigma_e^4 \frac{(4 - 8\rho + 4N\rho + 4\rho^2 - 3N\rho^2)}{(N-1)N(\rho-1)^2} \quad (2.7)$$

obtida por substituição de k por k_{otimo} de (2.6) em (2.4).

Pela Figura 2.6, observamos que a $Var(\widehat{\sigma}_a^2 | k_{otimo})$ é uma função crescente de ρ para qualquer N fixado⁴. Para cada valor de ρ , a variância é menor quanto maior for o valor de N . Além disto, verificamos que, para $\rho > 0,9$ a variância de $\widehat{\sigma}_a^2$ cresce rapidamente para infinito. Isso pode comprometer a estimação de σ_a^2 para $\rho > 0,9$.

Figura 2.6: Variância de $\widehat{\sigma}_a^2$ dado k_{otimo} para $0 < \rho < 0,8$ com diversos valores de N

Ainda que k_{otimo} possa assumir valores no intervalo $1 < k < N$, do ponto de vista prático, só pode assumir valores inteiros no intervalo $2 \leq k \leq N/2$. Neste caso, por (2.6), verificamos que

$$\frac{2}{2 - 2N + N^2} \leq \rho \leq \frac{N - 3}{2N - 3}.$$

⁴Este comentário também é válido quando a variância é descrita em função de θ (como na seção 2.1.1). Estamos comentando aqui, e não anteriormente, pois ρ tem domínio limitado e permite, assim, uma avaliação completa da $Var(\widehat{\sigma}_a^2 | k_{otimo})$.

Para $\rho \leq \frac{2}{2-2N+N^2}$ a melhor escolha de k é 2 e para $\rho > \frac{N-3}{2N-3}$, a melhor escolha de k é $\frac{N}{2}$. Para $\frac{2}{2-2N+N^2} \leq \rho \leq \frac{N-3}{2N-3}$, se k_{otimo} não for inteiro, dentre os dois divisores de N mais próximos de k_{otimo} , escolhemos aquele que produz a menor variância, como mostramos no exemplo da seção 2.1.1.

O programa PO.COM apresenta todos os possíveis planejamentos para N fixado pelo pesquisador. Então, mesmo desconhecendo o valor exato de ρ , o pesquisador pode escolher o planejamento ótimo procurando o provável intervalo a que ρ pertence.

2.2 Estimação de θ

O estimador de θ obtido pela ANOVA e apresentado na expressão (1.10) do capítulo anterior é um estimador viesado. Nesta seção, vamos desenvolver um estimador não-viesado para θ usando as propriedades da distribuição F de Fisher-Snedecor. A variância deste estimador será usada para o estudo de alocação ótima.

Teorema 4 (Searle, 1971) *Seja F_{v_1, v_2} uma variável com distribuição F central, com v_1 e v_2 graus de liberdade, então*

$$E(F_{v_1, v_2}) = \frac{v_2}{v_2 - 2}$$

e

$$Var(F_{v_1, v_2}) = \frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)}$$

Da seção 1.7.2, obtemos que $\mathcal{F} = \frac{QMA}{QMR}$ tem distribuição $(1 + n\theta)F_{k-1, N-k}$. Então:

$$E\left(\frac{QMA}{QMR}\right) = E[(1 + n\theta)F_{k-1, N-k}] = \frac{N - k}{N - k - 2}(1 + n\theta),$$

e, portanto, um estimador não-viesado para θ é:

$$\hat{\theta} = \frac{1}{n} \left[\frac{(N - k - 2)}{(N - k)} \frac{QMA}{QMR} - 1 \right]. \quad (2.8)$$

A variância deste estimador é calculada da seguinte forma:

$$Var(\hat{\theta}) = Var\left\{\frac{1}{n}\left[\frac{(N-k-2)QMA}{(N-k)QMR} - 1\right]\right\} = \frac{(N-k-2)^2}{n^2(N-k)^2} Var\left(\frac{QMA}{QMR}\right)$$

$$Var(\hat{\theta}) = \frac{(N-k-2)^2}{n^2(N-k)^2} Var\left[\left(1 + n\frac{\sigma_a^2}{\sigma_e^2}\right)F_{k-1, N-k}\right] = \frac{2(N-3)(1 + n\frac{\sigma_a^2}{\sigma_e^2})^2}{n^2(k-1)(N-k-4)}.$$

2.2.1 Variância de $\hat{\theta}$ expressa em função de θ

Vamos escrever $Var(\hat{\theta})$ desenvolvida anteriormente em função de N, k e θ . Então:

$$Var(\hat{\theta}) = \frac{2(N-3)}{(k-1)(N-k-4)}\left(\theta + \frac{k}{N}\right)^2, \quad 1 < k < N-4 \quad (2.9)$$

Para N fixado e θ conhecido, $Var(\hat{\theta})$ é uma função de k . Para facilitar o nosso estudo, vamos considerar que k varia continuamente no intervalo $(1, N-4)$. Como

$$\lim_{k \rightarrow 1} Var(\hat{\theta}) = \lim_{k \rightarrow N-4} Var(\hat{\theta}) = \infty,$$

e os termos que compõem $Var(\hat{\theta})$ são contínuos e positivos, $Var(\hat{\theta})$ tem pelo menos um mínimo no intervalo $1 < k < N-4$. A Figura 2.7 nos fornece uma melhor idéia sobre esta função.

Para determinarmos o ponto de mínimo, fazemos

$$\frac{dVar(\hat{\theta})}{dk} = 0$$

e obtemos a equação

$$k^2(2N\theta + N - 3) + 2k(N^2\theta^2 - N + 4) + (3N^2\theta^2 - N^3\theta^2 + 8N\theta - 2N^2\theta) = 0.$$

As raízes desta equação são:

$$k_1 = \frac{-2N^2\theta^2 - N^2\theta + 3N\theta}{2N\theta + N - 3} = -N\theta \quad e, \quad (2.10)$$

$$k_2 = \frac{N^2\theta - 3N\theta + 2N - 8}{2N\theta + N - 3} \quad (2.11)$$

Figura 2.7: $Var(\hat{\theta})$ como função de k para $N=40$ e θ fixo

Vamos estudar as raízes k_1 e k_2 como funções de θ . Lembramos que a raiz cujo contradomínio tem intersecção com o intervalo $(1, N-4)$ é o ponto que minimiza $Var(\hat{\theta})$.

i) Estudo de k_1

A raiz k_1 só assume valor negativo, isto é, seu contradomínio não contém o intervalo $(1, N-4)$. Portanto, não é o ponto de mínimo que procuramos.

ii) Estudo de k_2

Para N fixo, k_2 é uma função crescente de θ , tende a $\frac{2(N-4)}{N-3}$ quando θ tende a zero, e tende a $\frac{N-3}{2}$ quando θ tende a infinito (veja a Figura 2.8). Seu contradomínio está contido no intervalo $(1, N-4)$, portanto, a raiz k_2 é k_{otimo} que minimiza $Var(\hat{\theta})$.

Vimos na seção 2.1 que, somente quando θ está no intervalo $[\frac{2}{N(N-2)}, \frac{N-3}{N}]$, k_{otimo} obedece à restrição prática de pertencer ao intervalo $[2, \frac{N}{2}]$. Nesta seção, a situação é um pouco diferente. k_{otimo} , apesar de ser uma função crescente em θ , assume somente valores no intervalo $(2, \frac{N}{2})$. Portanto, k_{otimo} obedece à restrição prática, para qualquer θ .

Figura 2.8: k_{otimo} que minimiza $Var(\hat{\theta})$ como função de θ com diversos valores de N

Para determinarmos o planejamento com alocação ótima, substituímos θ por um dado valor (baseado nos experimentos anteriores ou intuição) na expressão (2.11). Se o valor de k_{otimo} obtido em (2.11) for divisor de N , o planejamento ótimo terá $k = k_{otimo}$ e $n = \frac{N}{k_{otimo}}$. Caso contrário, devemos escolher, entre os dois divisores de N mais próximos de k_{otimo} , aquele que produz a menor variância de $\hat{\theta}$.

Pela Figura 2.8, observamos que a função de k_{otimo} tende a se estabilizar para $\theta > 2$. Isto nos indica a possibilidade de adotarmos um mesmo planejamento a partir de $\theta = 2$. Um estudo mais cuidadoso dessa idéia intuitiva foi feito. Trabalhamos com $N=20, 40, 100, 200$ e 500 , usando o programa PO.COM, e vimos que, dependendo do tamanho da amostra, existe sempre um valor de θ a partir do qual, o planejamento $k = \frac{N}{2}$, $n=2$ sempre produz uma variância próxima à do planejamento com k_{otimo} . Este valor de θ é próximo de 2 e, diminui quando o tamanho da amostra aumenta, por exemplo: para $N=20$, a partir de $\theta = 2,12$ o planejamento $k = N/2$ deve ser adotado; e para $N = 500$, este valor de θ diminui para 1,38.

Concluimos então que, quando $\theta > 2$, o planejamento $k = \frac{N}{2}$, $n=2$ é sempre uma boa opção. Para $\theta < 2$, substituímos o valor de θ em (2.11) para obter o k_{otimo} . Quando este não é divisor

de N , escolhemos dentre os dois divisores de N mais próximos de k_{otimo} , aquele que produz a menor variância.

No caso em que θ é desconhecido, o programa PO.COM apresenta todos os possíveis planejamentos para o tamanho da amostra N pré-definido pelo pesquisador, fornecendo uma visão completa sobre os planejamentos e a variância do estimador $\hat{\theta}$ em cada planejamento.

2.2.2 Variância de $\hat{\theta}$ expressa em função de ρ

Nesta subseção, vamos estudar o planejamento ótimo na estimação de θ , escrevendo a variância do estimador em função de ρ . O estimador usado é o estimador não-viesado apresentado no início da seção 2.2. Seguiremos os mesmos procedimentos das seções anteriores.

Para escrever $Var(\hat{\theta})$ como função de ρ substituímos θ por $\frac{\rho}{1-\rho}$ em (2.9), obtendo:

$$Var(\hat{\theta}) = \frac{2(N-3)[N\rho + k(1-\rho)]^2}{N^2(k-1)(N-k-4)(1-\rho)^2}, \quad 1 < k < N-4 \text{ e } \rho \neq 1 \quad (2.12)$$

Para $1 < k < N-4$, $Var(\hat{\theta})$ tende a infinito quando k tende a 1 ou $(N-4)$. Existe pelo menos um ponto que minimiza $Var(\hat{\theta})$ neste intervalo. Este ponto de mínimo é a raiz da equação da primeira derivada de $Var(\hat{\theta})$ em relação a k , cujo contradomínio apresenta intersecção com o intervalo $(1, N-4)$. Repetindo o mesmo procedimento das seções anteriores, obtemos

$$k_1 = \frac{-N\rho}{1-\rho} \text{ e,} \quad (2.13)$$

$$k_2 = \frac{N^2\rho - 5N\rho + 2N - 8 + 8\rho}{N\rho + N - 3 + 3\rho} \quad (2.14)$$

Vamos estudar as raízes k_1 e k_2 como funções de ρ .

i) Estudo de k_1

O contradomínio da raiz k_1 só assume valor negativo. Não tem intersecção com o intervalo $(1, N-4)$. Portanto, não é a raiz que conduz a uma variância mínima.

ii) Estudo de k_2

A raiz k_2 é uma função crescente de ρ , como podemos constatar com a Figura 2.9. k_2 tende a $\frac{2(N-4)}{N-3}$ quando ρ tende a zero e tende a $\frac{N-3}{2}$ quando ρ tende a um. Seu contradomínio está contido no intervalo $(1, N-4)$, quando $\rho \in [0, 1]$. É o k_{otimo} que procuramos.

Figura 2.9: k_{otimo} que minimiza $Var(\hat{\theta})$ como função de ρ com diversos valores de N

Assim como nas seções anteriores, para obtermos o valor de k_{otimo} que minimiza $Var(\hat{\theta})$ em uma amostra de tamanho N , devemos substituir o valor de ρ em (2.14) por estimativas obtidas de experimentos anteriores, ou por intuição.

A variância do estimador de θ dado k igual a k_{otimo} , obtida de (2.12) é:

$$Var(\hat{\theta}|k_{otimo}) = \frac{8(N-3)(N\rho - \rho + 1)(N-4+4\rho)}{N^2(1-\rho)^2(N-5)^2} \quad (2.15)$$

Através da Figura 2.10, observamos que $Var(\hat{\theta}|k_{otimo})$ é uma função crescente de ρ para qualquer N . Além disto, para cada valor de ρ , a variância é menor quanto maior for o valor de N , e para os valores de ρ maiores do que 0,9, $Var(\hat{\theta}|k_{otimo})$ cresce rapidamente para o infinito.

Figura 2.10: Variância de $\hat{\theta}$ dado k_{otimo} para $0 < \rho < 0,8$ com diversos valores de N

Isto pode comprometer o resultado da estimação de θ para $\rho > 0,9$.

Assim como observado na seção 2.2.1, o planejamento fixo com $k = \frac{N}{2}$ e $n = 2$ passa a ser a opção ideal a partir de um certo valor de ρ (digamos ρ^*). Este valor de ρ é menor quanto maior o tamanho da amostra.

Os outros planejamentos de uma amostra com tamanho N, para os quais $\rho < \rho^*$, podem ser obtidos através do programa PO.COM.

2.3 Uso do programa PO.COM

Mencionamos na seção 2.1.1 que, a motivação inicial para a elaboração do programa PO.COM é o de contornar os problemas de desconhecimento ou a falta de estimativa confiável do parâmetro θ (ou ρ) na determinação do valor de k_{otimo} . Porém, durante o desenvolvimento deste estudo, percebemos que, mesmo conhecendo o valor de θ (ou ρ), k_{otimo} obtido, muitas vezes, não pode ser utilizado por não ser divisor de N; neste caso, K SUGERIDO apresentado

na saída do programa pode ser útil - ele é escolhido entre os dois divisores mais próximos de k_{otimo} por produzir a menor variância.

Para acessar o programa, digite primeiro PO.COM, e siga as instruções que o sistema fornece, alimentando-o com as informações básicas: tamanho da amostra e o parâmetro a ser estimado. O programa fornece todos os possíveis planejamentos para aquele tamanho da amostra como apresentado na Tabela 2.1.

Tabela 2.1: Arquivo de saída do programa PO.COM

O parametro a ser estimado THETA

O tamanho da amostra: 200

Os divisores de N sao: 2 4 5 8 10 20 25 40 50 100

THETA	RHO	KOTIMO	K SUGERIDO	VARIANCIA DO KOTIMO	VARIANCIA DO K
0.01	0.01	3.93	4	0.0006	0.0006
0.02	0.02	5.83	5	0.0011	0.0011
0.03	0.03	7.69	8	0.0015	0.0015
0.04	0.04	9.52	10	0.0020	0.0020
0.08	0.07	14.78	20	0.0035	0.0036
0.14	0.12	22.92	25	0.0065	0.0066
0.23	0.19	33.13	40	0.0121	0.0122
0.41	0.29	45.74	50	0.0238	0.0239
1.44	0.59	73.89	100	0.1448	0.1559

Cada linha da Tabela 2.1 representa um planejamento em que $k = K \text{ SUGERIDO}$ e $n = \frac{N}{k}$.

As colunas de THETA e RHO apresentam os valores de θ e de ρ , a partir dos quais, se inicia um novo planejamento. K SUGERIDO é o divisor de N que produz a variância mais próxima daquela obtida com k_{otimo} para o valor de θ (ou ρ) apresentado no início da linha. VARIANCIA DO K é a variância do estimador quando K SUGERIDO é adotado para aquele valor de θ (ou ρ) fixado.

Os planejamentos apresentados na Tabela 2.1 atende todos os possíveis valores de θ (ou ρ).

Para aqueles valores de θ (ou ρ) não encontrados na tabela, basta localizar na coluna THETA (ou RHO) o intervalo a que ele pertence e adotar o K SUGERIDO na linha de onde inicia o intervalo. Por exemplo: Em um experimento se $N=200$ e $\theta = 0,07$, verificamos na Tabela 2.1 que $0,04 < \theta < 0,08$, então $k = 10$ classes sugerido pelo programa é o planejamento ótimo que nos fornece a menor variância. Se $\theta = 2$, o planejamento ótimo será escolher $k = 100$ classes e, conseqüentemente, $n = 2$ unidades experimentais em cada classe.

Se não conhecemos o valor de θ (ou ρ), porém sabemos que ele pode estar em um determinado intervalo e, o programa pode sugerir vários possíveis planejamentos, então todos os planejamentos poderão ser utilizados, no entanto, aquele que fornece maior agilidade ou viabilidade prática é sempre recomendado. Por exemplo: Desejamos realizar um experimento com 200 unidades experimentais. Sabemos que $0,2 < \rho < 0,4$. Pela Tabela 2.1 vemos que, os planejamentos com $k=40$ e $k=50$ podem ser adotados. Suponhamos que o custo de aumentar o número de unidades experimentais em uma classe é mais baixo do que aumentar o número de classes. Escolhemos então o planejamento com $k=40$, que tem número de classes menor para reduzir o custo total do experimento.

Capítulo 3

Planejamento Ótimo na Estimação de

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

No capítulo anterior estudamos o planejamento com alocação ótima na estimação dos parâmetros σ_a^2 e θ do modelo de efeito aleatório com um fator. Neste capítulo, vamos estudar o planejamento ótimo na estimação de outro parâmetro importante deste modelo, aliás o parâmetro de nosso principal interesse - o coeficiente de correlação intraclassa (ρ) - definido pela expressão (1.4) no Capítulo 1.

Introduzido por Fisher(1925), ρ reflete a variação relativa entre a variância das classes e a variância total das observações do modelo. É uma medida descritiva simples, mas tem ampla aplicação em diferentes áreas. Em genética, por exemplo, ρ é utilizado para estimar a herdabilidade de alguma característica na população animal ou vegetal; em Psicologia é uma medida importante no estudo de confiabilidade; em Epidemiologia mede o grau de semelhança familiar por fator biológico ou característica ambiental; em Análise de Sensibilidade ρ é usado para medir a eficiência de um tratamento experimental.

O estudo inferencial sobre ρ foi bastante desenvolvido nessas últimas décadas. Donner(1986) acompanhou esse desenvolvimento, mostrando as teorias recentes de estimação e testes de hipóteses. Entre os diversos métodos de estimação de ρ citados na literatura, o método de

análise de variância e o de máxima verossimilhança, são os mais usados, e fornecem estimadores mais eficientes quando comparados com outros métodos. Por essa razão, foram os escolhidos para o nosso estudo de planejamento ótimo. Nas seções subseqüentes, vamos fazer um estudo similar ao do capítulo anterior para cada estimador obtido pelos métodos citados acima.

3.1 Estimador de Análise de Variância

Como mostramos no Capítulo 1, o estimador de ρ é obtido por substituição dos estimadores dos componentes de variância obtidos pela Análise de Variância na expressão (1.4). Temos então

$$\hat{\rho}_A = \frac{(QMA - QMR)}{QMA + (n_0 - 1)QMR}, \quad (3.1)$$

onde $n_0 = n$, se os dados forem balanceados, isto é, se o número de observações em cada classe forem iguais (para os dados não-balanceados, $n_0 = [N - \sum_{i=1}^k n_i^2/N]/(k-1)$).

Esse estimador é viesado, porém consistente. O problema, contudo, mais sério desse estimador é que ele pode fornecer uma estimativa negativa, o que contradiz a definição de ρ que tem que ser positivo. Segundo Scheffé(1959), quando isso acontecer, essa estimativa deve ser substituída pelo valor zero.

A variância do estimador de ρ , determinada por Smith(1956), sob a suposição de normalidade, com dados não-balanceados, para amostra de tamanho grande é dado por

$$Var(\hat{\rho}_A) = \frac{2(1-\rho)^2}{n_0^2} \left\{ \frac{[1 + \rho(n_0 - 1)]^2}{N - k} + \frac{(k-1)(1-\rho)[1 + \rho(2n_0 - 1)]}{(k-1)^2} + \frac{\rho^2[\sum n_i^2 - 2N^{-1} \sum n_i^3 + N^{-2}(\sum n_i^2)^2]}{(k-1)^2} \right\}$$

No caso de dados balanceados, a variância é reduzida a

$$Var(\hat{\rho}_A) = \frac{2(1-\rho)^2[k + (N-k)\rho]^2}{N(N-k)(k-1)}, \quad 1 < k < N \quad (3.2)$$

Como os termos que compõem $Var(\hat{\rho}_A)$ são positivos e contínuos, e $Var(\hat{\rho}_A)$ tende para infinito quando k tende aos seus extremos 1 e N , existe pelo menos um ponto no intervalo $(1, N)$ que minimiza $Var(\hat{\rho}_A)$. (Veja a Figura 3.1)

Figura 3.1: Variância de $\hat{\rho}_A$ como função de k para diversos valores de ρ ($N=40$)

Fazendo $dVar(\hat{\rho}_A)/dk = 0$, temos

$$2(1 - \rho)(N - k)(k - 1) - [k + (N - k)\rho](N - 2k + 1) = 0$$

Resolvendo essa equação, obtemos

$$k_A = \frac{N[(N - 1)\rho + 2]}{[(N + 1) + (N - 1)\rho]} \quad (3.3)$$

Para N fixo, k_A é uma função crescente de ρ (veja a Figura 3.2). Tende a $\frac{2N}{N+1}$ quando ρ tende a zero, e tende a $\frac{N+1}{2}$ quando ρ tende a um. O contradomínio de k_A está contido no intervalo $(1, N)$. Portanto, k_A é k_{otimo} que minimiza a $Var(\hat{\rho}_A)$.

Como no capítulo anterior, esta solução depende do valor de ρ , que é geralmente desconhecido. Neste caso, substituímos ρ por uma estimativa obtida em experimentos anteriores ou baseada na intuição. Portanto, se o pesquisador dispuser desta informação, o problema está resolvido - precisando apenas ajustar o k_A de tal forma que ele seja divisor de N como vimos no capítulo anterior. No caso destas informações preliminares não estarem disponíveis, será que é possível determinarmos um planejamento "quase-ótimo", cuja variância ao estimar ρ seja

Figura 3.2: k_A como função de ρ para diversos valores de N

próxima à da variância ótima para qualquer valor de ρ ? Em outras palavras, é possível determinar um planejamento que não depende do valor de ρ ? Vamos realizar um estudo baseado na variância ótima (como função de ρ) para encontrar este planejamento.

a) Estudo de $Var(\hat{\rho}_A|k_A)$

A variância de $\hat{\rho}_A$ com a adoção de k_A é obtida por substituição de (3.3) em (3.2). Então

$$Var(\hat{\rho}_A|k_A) = \frac{8(1-\rho)^2[(N-1)\rho+1]}{(N-1)^2}, \quad (3.4)$$

A Figura 3.3 mostra a função apresentada em (3.4) para diversos tamanhos de amostra.

Para $0 < \rho < 1$, $Var(\hat{\rho}_A|k_A)$ atinge seu valor máximo no ponto

$$\rho = \frac{(N-3)}{3(N-1)} \quad (3.5)$$

e o valor correspondente de k_A neste ponto é

$$k_A = \frac{N+3}{4} \quad (3.6)$$

b) Determinação do Planejamento Quase-Ótimo

Figura 3.3: Variância ótima como função de ρ para diversos valores de N

Para encontrar um único planejamento que se aproxima do planejamento ótimo para $0 < \rho < 1$, vamos considerar um planejamento que se aproxima do planejamento ótimo quando $\rho = \frac{N-3}{3(N-1)}$ e verificar o seu comportamento para outros valores de ρ .

Vimos antes que quando $\rho = \frac{N-3}{3(N-1)}$, $k = \frac{N+3}{4}$ é o planejamento ótimo. Para N suficientemente grande, a variância do estimador no planejamento $k = N/4$ coincide com a variância do planejamento ótimo quando $\rho = \frac{N-3}{3(N-1)}$. Vamos estudar o comportamento da variância deste planejamento para outros valores de ρ .

Adotando o planejamento $k = N/4$ e $n = 4$, temos que a variância de $\hat{\rho}_A$ é:

$$Var(\hat{\rho}_A|k = \frac{N}{4}) = \frac{2(1-\rho)^2(1+3\rho)^2}{3(N-4)} \quad (3.7)$$

A Figura 3.4 mostra, para $N = 40$, que a $Var(\hat{\rho}_A|k = N/4)$ é visualmente muito próxima de $Var(\hat{\rho}_A|k = k_A)$ para $\rho > 0,2$. Para comparar analiticamente os dois planejamentos, construímos a função $G_A(\rho) = Var(\hat{\rho}_A|k = N/4) - Var(\hat{\rho}_A|k = k_A)$ (veja a Figura 3.5).

$G_A(\rho)$ é uma função contínua com as seguintes propriedades:

- i) - $G_A(0) = 2(N-7)^2/3(N-4)(N-1)^2$ e $G_A(1) = 0$

Figura 3.4: Variância de $\hat{\rho}_A$ para $k = k_A$ e $k = N/4$ ($N = 40$)

Figura 3.5: A função diferença $G_A(\rho)$

- ii) - para $0 < \rho < 1$, G_A atinge o valor mínimo em $\rho = (N-7)/3(N-1)$, com $G_A(\frac{N-7}{3(N-1)}) = 0$
- iii) - para $\rho > \frac{N-7}{3(N-1)}$, G_A atinge o valor máximo em $\rho = (2N-5)/3(N-1)$, com $G_A(\frac{2N-5}{3(N-1)}) = 2(N+2)^4/27(N-4)(N-1)^4$.

Com as propriedades de $G_A(\rho)$ apresentadas acima, observamos que, para $\rho > \frac{N-7}{3(N-1)}$, a variância do estimador na adoção do planejamento com $k = \frac{N}{4}$ é próxima daquela obtida com $k = k_A$. Ainda que a diferença entre os dois planejamentos seja máxima, pela propriedade iii), podemos observar que se N for suficientemente grande, essa diferença pode ser ignorada. Isto é, as duas funções tornam-se mais próximas à medida que N cresce.

Assim como no capítulo anterior, se temos alguma informação acerca do valor de ρ , podemos determinar o planejamento ótimo. O programa PO.COM está disponível também para o fornecimento do planejamento ótimo. Contudo, $k = N/4$ é o único planejamento, cuja variância do estimador se aproxima daquela obtida com k_{otimo} , em quase todo o intervalo de variação de ρ . Portanto, em caso de desconhecimento total sobre o valor de ρ , o planejamento com $k = \frac{N}{4}$, é sempre uma boa opção na estimação de ρ . Para utilizarmos este planejamento, na prática, devemos ajustar N de forma que ele seja múltiplo de 4.

3.2 Estimador de Máxima Verossimilhança

Quando os dados são não-balanceados, não existe expressão explícita para o estimador de máxima verossimilhança ($\hat{\rho}_M$), porém a solução numérica pode ser obtida através de um algoritmo iterativo. Para dados balanceados, $\hat{\rho}_M$ é reduzido a:

$$\hat{\rho}_M = [(k-1)QMA - kQMR]/[(k-1)QMA + k(n-1)QMR]. \quad (3.8)$$

Segundo Donner & Koval(1980), a variância de $\hat{\rho}_M$ para amostras grandes, no caso geral, é expressa explicitamente da seguinte forma:

$$Var(\hat{\rho}_M) = \frac{2N(1-\rho)^2}{N \sum n_i(n_i-1)U_i U_i^{-2} - \rho^2 [\sum n_i(n_i-1)U_i^{-1}]^2}$$

onde $U_i = 1 + (n_i - 1)\rho$ e $V_i = 1 + (n_i - 1)\rho^2$

Para dados balanceados, a variância para amostra grande é reduzida a

$$Var(\hat{\rho}_M) = \frac{2(1 - \rho)^2[k + (N - k)\rho]^2}{N(N - k)k}, \quad 0 < k < N \quad (3.9)$$

Como os limites de $Var(\hat{\rho}_M)$ tendem a infinito nos extremos 0 e N, e os termos que compõem $Var(\hat{\rho}_M)$ são contínuos e positivos, existe pelo menos um ponto no intervalo $(0, N)$ que minimiza $Var(\hat{\rho}_M)$. (Veja a Figura 3.6)

Figura 3.6: Variância de $\hat{\rho}_M$ como função de k para diversos valores de ρ (N=40)

Seguindo o mesmo raciocínio das seções anteriores, o ponto de mínimo é obtido, fazendo $dVar(\hat{\rho}_M)/dk = 0$, isto é,

$$2(1 - \rho)(N - k)k - [k + (N - k)\rho](N - 2k) = 0, \quad e$$

resolvendo esta equação, obtemos

$$k_M = \frac{N\rho}{1 + \rho} \quad (3.10)$$

Para $0 < \rho < 1$, o contradomínio de k_M está contido no intervalo $(0, N)$, portanto k_M é o k_{otimo} que minimiza $Var(\hat{\rho}_M)$. A Figura 3.7 mostra k_M em função de ρ .

Figura 3.7: k_M como função de ρ para diversos valores de N

Na seção anterior, determinamos um planejamento quase-ótimo, em que a variância do estimador é sempre próxima daquela obtida com k_{otimo} para todos os valores de ρ . Nesta seção, vamos realizar o mesmo estudo.

Para estudar a variância do estimador com a adoção de k_{otimo} , substituímos k_M em (3.9). Obtemos:

$$Var(\hat{\rho}_M | k_M) = 8(1 - \rho)^2 \rho / N, \quad (3.11)$$

e o seu comportamento pode ser visto na Figura 3.8.

$Var(\hat{\rho}_M | k_M)$ atinge o máximo quando $\rho = 1/3$. O valor correspondente de k_M neste ponto é

$$k_M = \frac{N}{4}.$$

Será que este é o planejamento quase-ótimo que procuramos? A variância de $\hat{\rho}_M$ quando adotamos este planejamento é:

$$Var(\hat{\rho}_M | k = \frac{N}{4}) = \frac{2(1 - \rho)^2(1 + 3\rho)^2}{3N} \quad (3.12)$$

Pela Figura 3.9, vemos que, a variância de $\hat{\rho}_M$ no planejamento com $k = N/4$ é próxima

Figura 3.8: Variância ótima como função de ρ para diversos valores de N

da variância com $k = k_{otimo}$ para $0 < \rho < 1$. Para verificar isso analiticamente, construímos a função de diferença $G_M(\rho) = Var(\hat{\rho}_M|k = N/4) - Var(\hat{\rho}_M|k = k_{otimo})$, cujo comportamento está representado na Figura 3.10. As propriedades desta função são:

- i) - $G_M(0) = 2/3N$ e $G_M(1) = 0$
- ii) - atinge o valor mínimo em $\rho = \frac{1}{3}$, e $G_M(\frac{1}{3}) = 0$
- iii) - para $\rho > \frac{1}{3}$, G_M atinge o valor máximo em $\rho = \frac{2}{3}$ e $G_M(\frac{2}{3}) = 2/27N$.

Observando as propriedades, vemos que o planejamento proposto com $k = N/4$ é muito próximo do planejamento com k_{otimo} para $\rho > 0,2$. Ainda que a diferença entre os dois planejamentos seja máxima, se N é grande, então essa diferença torna-se desprezível.

Concluimos então, que o planejamento $k = N/4$ é o planejamento quase-ótimo, e pode ser utilizado qualquer que seja o valor de $\rho > 0,2$. Para valores de $\rho < 0,2$ devemos utilizar planejamentos com valores de k menores do que $k/4$, que serão determinados com o auxílio do programa PO.COM.

Figura 3.9: Variância de $\hat{\rho}_M$ para $k = k_M$ e $k = N/4$ ($N=40$)

Figura 3.10: A função diferença $G_M(\rho)$

Capítulo 4

Coeficiente de Correlação Intraclass como Medida de Confiabilidade

" The most elegant design of a clinical study will not overcome the damage caused by unreliable or imprecise measurement. " - Fleiss

A qualidade de uma pesquisa é garantida, antes de tudo, pela qualidade dos dados. Se em uma pesquisa, os dados utilizados na análise são obtidos de maneira imprecisa, seja por falha humana, ou má-calibração dos equipamentos utilizados, a validação dos resultados baseada em tal conjunto de dados é duvidosa. Esta imprecisão pode levar, por exemplo, a uma conclusão errada ou aumento de custos.

Sendo a confiabilidade um sinônimo de qualidade dos dados, parece fundamental que o pesquisador avalie a confiabilidade de seus dados antes da realização de um estudo maior. Segundo Fleiss (1986), para processar esta avaliação, basta que se analise uma amostra de 15 a 20 indivíduos nos casos em que as variáveis sejam quantitativas, para variáveis qualitativas a amostra deve ser maior.

É importante lembrarmos ainda que, a confiabilidade deve ser avaliada toda vez que um experimento é realizado, pois a alta confiabilidade em um estudo não implica necessariamente em alta confiabilidade em estudos posteriores.

Para cada tipo de dados -quantitativos ou qualitativos- existem medidas de confiabilidade apropriadas. Como exemplo podemos citar o coeficiente de correlação intraclasse que é uma boa medida para dados quantitativos, e o coeficiente Kappa que é uma boa medida para dados qualitativos. Neste capítulo, vamos apresentar a aplicação do coeficiente de correlação intraclasse no estudo de confiabilidade. Outras medidas de confiabilidade podem ser encontradas em Bartko(1976), por exemplo.

Iniciamos a seção 4.1, introduzindo o conceito de confiabilidade. Relacionamos este conceito com o coeficiente de correlação intraclasse (ρ) definido através do modelo de efeitos aleatórios com um fator. Em seguida, na seção 4.2, apresentamos a estimação pontual e por intervalo de ρ , e estudamos as situações em que os resultados das duas formas de estimação são similares. Na seção 4.3, apresentamos algumas conseqüências da não-confiabilidade quantificadas como funções de ρ , e mostramos como o grau de confiabilidade pode interferir em um estudo. Na seção 4.4, mostramos como controlar a confiabilidade através do número de réplicas nas classes. Na seção 4.5, determinamos o número de réplicas necessário para a minimização do custo. Na seção 4.6, adicionando um outro fator ao modelo, estudamos ρ e sua interpretação nos modelos com dois fatores (aleatório e misto). Na última seção, um exemplo real é apresentado. Este exemplo é um estudo sobre o nível sérico de colesterol (concentração de colesterol no sangue) realizado no Hospital Universitário da Universidade de São Paulo. Através deste exemplo, podemos ver a aplicação de alguns resultados deste capítulo.

4.1 Confiabilidade e ρ

Seja y_i o valor observado de uma determinada variável em um indivíduo i . Não importa a que y_i se refere, nem como y_i foi obtido. Falamos da não-confiabilidade do valor y_i se, ao repetirmos a observação várias vezes sob condições similares, os valores de y_i forem diferentes entre si. Seja T_i o valor verdadeiro da variável para um indivíduo (em psicometria, T_i é dita “escore verdadeiro” ou “escore livre de erro”). y_i pode diferir de T_i por uma série de causas, por exemplo, pela dificuldade de leitura no equipamento utilizado, pelo mau entendimento da

pergunta pelo entrevistado, ou pelo mau entendimento da resposta pelo entrevistador, etc. Estas causas são denominadas erro aleatório, e as denotamos por ϵ_i . Podemos então escrever:

$$Y_i = T_i + \epsilon_i$$

Para avaliarmos este erro aleatório associado a cada observação, sugerimos um experimento que consiste em sortear uma amostra de k indivíduos e avalia cada indivíduo um certo número de vezes (n_i repetições para o i -ésimo indivíduo). Os resultados deste experimento podem ser analisados por um modelo de efeitos aleatórios com um fator considerado anteriormente em (1.1). Isto é,

$$Y_{ij} = \mu + A_i + \epsilon_{ij} \quad i = 1, \dots, k; \quad j = 1, \dots, n_i;$$

onde

k : número de indivíduos;

n_i : número de observações (réplicas) do i -ésimo indivíduo.

Cabe lembrar que $A_i \sim Normal(0, \sigma_a^2)$, $\epsilon_{ij} \sim Normal(0, \sigma_e^2)$ e são completamente independentes.

O coeficiente de correlação intraclassa que mede a confiabilidade dos dados é:

$$\rho = \frac{\sigma_a^2}{\sigma_y^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

Notemos que ρ é o quociente entre duas variâncias, e está limitado entre 0 e 1. ρ é uma medida relativa que descreve a “gravidade” das variações.

Dizemos que a confiabilidade é alta, se o valor assumido por ρ for grande, isto é, quando a variação das medidas para cada indivíduo for pequena em relação à variação das medidas entre os indivíduos, caso contrário, se ρ assumir um valor baixo, a confiabilidade é baixa. Escrevendo ρ em função de $\theta = \frac{\sigma_a^2}{\sigma_e^2}$, torna-se fácil verificar estas afirmações.

Vimos no Capítulo 1 que ρ é a correlação entre as observações de uma mesma classe. Aqui, cada indivíduo é considerado como uma classe, e é fácil observar que quanto maior for

a correlação entre as réplicas, menor é a variabilidade entre as réplicas e, portanto, maior será a confiabilidade.

Em geral, não existe uma regra rígida para determinarmos quando a confiabilidade de uma medida é alta ou baixa. Essa definição depende muito da área e do estudo que está sendo realizado. Segundo Fleiss(1986), em geral, quando ρ é menor que 0,4, as medidas têm confiabilidade baixa, se ρ estiver entre 0,4 e 0,75, a confiabilidade nas medidas é razoável, e se ρ for maior que 0,75, então as medidas terão excelente confiabilidade.

4.2 Estimação da confiabilidade de uma medida

Para avaliarmos a confiabilidade de uma medida, várias réplicas de uma mesma classe são obtidas. Essa classe pode ser indivíduo, animal, etc. Por exemplo, queremos avaliar a confiabilidade das medidas da atividade de uma certa enzima em um laboratório. Para isso, uma amostra de sangue de um paciente é coletada e dividida em várias porções para serem analisadas no laboratório. Então, cada amostra de sangue (indivíduo) é considerada como uma classe, e os resultados das porções, réplicas. Outro exemplo: sabendo que a precisão da mensuração da taxa de colesterol em laboratórios é alta, queremos conhecer a variação natural da taxa de colesterol do ser humano. Então várias amostras de sangue de vários indivíduos são coletadas em tempos diferentes. Aqui, cada indivíduo representa uma classe, e as amostras de sangue em tempos diferentes são réplicas. Nos exemplos, a ordem de cada réplica não é considerada no estudo, isto é, não importa qual réplica foi obtida primeiro, o que importa é se todas foram obtidas em condições similares.

Neste tipo de estudo, os dados podem ser apresentados na forma da Tabela 4.1, e a estrutura destes dados nos permite ajustar o modelo com 1 fator aleatório.

Através do método de ANOVA apresentado no Capítulo 1, podemos obter a estimativa de ρ , a qual será denotada por R , como segue:

$$R = \frac{\widehat{\sigma_a^2}}{\widehat{\sigma_a^2} + \widehat{\sigma_e^2}} = \frac{QMA - QMR}{QMA + (n_0 - 1)QMR} \quad (4.1)$$

Tabela 4.1: Estrutura de dados para um estudo balanceado com k indivíduos

Indivíduos	Réplicas					
	1	2	..	j	..	n
1	X_{11}	X_{12}	..	X_{1j}	..	X_{1n}
2	X_{21}	X_{22}	..	X_{2j}	..	X_{2n}
3	X_{31}	X_{32}	..	X_{3j}	..	X_{3n}
.	...					
.	...					
i	X_{i1}	X_{i2}	..	X_{ij}	..	X_{in}
.	...					
.	...					
k	X_{k1}	X_{k2}	..	X_{kj}	..	X_{kn}

onde

$$n_0 = \begin{cases} n, & \text{se todos os indivíduos têm o mesmo número de réplicas;} \\ [N - \sum_{i=1}^k n_i^2/N]/(k-1) & \text{caso contrário.} \end{cases}$$

Se as distribuições dos efeitos A_i 's e dos erros aleatórios ϵ_{ij} 's são supostamente normais, também podemos construir intervalos de confiança para ρ . Se $n_0 = n$, os intervalos de confiança construídos para ρ serão exatos; caso contrário, se os números de réplicas para cada indivíduo forem diferentes, o intervalo será aproximado (Ver Capítulo 1 para maiores detalhes). O limite inferior deste intervalo é dado pela seguinte expressão:

$$\rho \geq \frac{\frac{QMA}{QMR} - F_{k-1, N-k, \alpha}}{\frac{QMA}{QMR} + (n_0 - 1)F_{k-1, N-k, \alpha}} = R_I \quad (4.2)$$

onde $F_{k-1, N-k, \alpha}$ é obtido da tabela da distribuição F, com $k-1$ e $N-k$ graus de liberdade, e probabilidade α na cauda superior.

A construção de intervalos para ρ serve para dois propósitos. O primeiro deles é testar a hipótese $H_0 : \rho = 0$. Quando o valor zero pertence ao intervalo construído, H_0 não é rejeitada, significando que as medidas possuem um grau de confiabilidade muito baixo, e que a maior parte da variação da medida é devido ao erro aleatório. Em situações práticas, este critério para avaliar a confiabilidade dos dados não é muito razoável, pois dependendo da área e/ou do estudo realizado podemos sempre rejeitar H_0 , o que não implica necessariamente em alta confiabilidade dos dados.

O segundo propósito e de maior importância é indicar o limite de incerteza apresentado pelo conjunto de dados submetido à análise. Este limite de incerteza determina o menor valor que ρ pode assumir. Isso nos permite visualizar o campo de variação de ρ . Apesar de, neste caso, termos maior informação sobre ρ , ainda não conseguimos dizer com certeza o grau de confiabilidade dos dados. Para fazermos afirmações sobre a confiabilidade dos dados, lembramos, mais uma vez, que é fundamental conhecer a área e/ou o estudo que está sendo realizado.

Para tornar mais claro o que foi exposto acima, consideremos o seguinte exemplo:

Exemplo: A tabela abaixo apresenta os resultados de um estudo de confiabilidade, usando "Hamilton's (1960) rating scale" para depressão.

pacientes	n_i	$\bar{X}_i$	s_i^2
1	3	9,6667	2,3333
2	2	20,5000	12,5000
3	2	13,5000	4,5000
4	4	11,7500	4,2500
5	3	22,3333	2,3333
6	3	14,0000	7,0000
7	4	18,2500	11,5833
8	2	11,0000	8,0000

Para estes dados, construímos a seguinte tabela de ANOVA:

f.v.	g.l.	QM
Entre indivíduos	7	61,5390
Dentre indivíduos	15	6,3889

Da tabela acima, obtemos $R = 0,75$ e $n_0 = 2,845$.

Usando (4.2) para $\alpha = 0,05$, obtivemos $R_I = 0,47$. Então o intervalo de confiança para ρ com 95 % de confiança é $\rho \geq 0,47$. Este intervalo é aproximado, pois os n_i 's são distintos.

Notemos que, R_I não é necessariamente próximo da confiabilidade estimada R . De fato, como vimos no exemplo anterior, a estimativa pontual para ρ foi 0,75, enquanto o limite de incerteza R_I obtido foi 0,47.

Qual a razão para essa diferença entre R e R_I ? Em que situações estão próximos ou afastados um do outro? A Tabela 4.2 mostra esta relação em várias tabelas de ANOVA, com graus de liberdade e QMR fixos, variando-se apenas $\mathcal{F}$ - a razão entre QMA e QMR. Os valores das colunas de R e R_I são calculados através das expressões (4.1) e (4.2).

Tabela 4.2: Estudo de R e R_I (variando $\mathcal{F}$)

	g.l.	QM	$F_{n1,n2,5\%}$	n	$\mathcal{F}$	R	R_I
Entre	9	3,5192	3,02	2	451	0,99557	0,986701
Dentro	10	0,0078					
Entre	9	2,34	3,02	2	300	0,993355	0,980067
Dentro	10	0,0078					
Entre	9	0,78	3,02	2	100	0,980198	0,941370
Dentro	10	0,0078					
Entre	9	0,39	3,02	2	50	0,960784	0,886080
Dentro	10	0,0078					
Entre	9	0,156	3,02	2	20	0,904761	0,737619
Dentro	10	0,0078					
Entre	9	0,039	3,02	2	5	0,666666	0,246882
Dentro	10	0,0078					
Entre	9	0,0078	3,02	2	1	0	-0,50248
Dentro	10	0,0078					

Notamos que, na Tabela 4.2, a distância entre a confiabilidade estimada R e o limite inferior

(R_I) está intimamente ligada a $\mathcal{F}$. Quanto maior for $\mathcal{F}$, mais próximos entre si estarão R e R_I . Este é um fato que ocorre em todas situações? Isto é, a distância entre R e R_I depende só de $\mathcal{F}$? e quanto aos graus de liberdade? Para verificarmos isso, escrevemos (4.1) e (4.2) como funções de $\mathcal{F}$. É fácil de observar que, quando $\mathcal{F}$ tende ao infinito, R e R_I tendem a um, caso contrário, a distância entre os dois valores dependem do valor da $F_{k-1, N-k, p}$ obtido pela tabela da distribuição Fisher-Snedecor, que é determinado pelos graus de liberdade. Consideremos agora um outro estudo, onde $\mathcal{F}$ é fixo e vamos variar apenas os graus de liberdade. Os resultados estão sendo exibidos na Tabela 4.3.

Tabela 4.3: Estudo de R e R_I (variando graus de liberdade)

	g.l.	QM	$F_{n1, n2, 5\%}$	n	$\mathcal{F}$	R	R_I
Entre	9	3,5192	2,39	2	451	0,9933801	0,984274
Dentro	20	0,0078					
Entre	9	3,5192	2,21	4	451	0,991192	0,980690
Dentro	30	0,0078					
Entre	9	3,5192	2,12	5	451	0,989015	0,976939
Dentro	40	0,0078					
Entre	9	3,5192	1,94	20	451	0,957463	0,920498
Dentro	190	0,0078					
Entre	10	3,5192	2,85	2	451	0,995576	0,987445
Dentro	11	0,0078					
Entre	20	3,5192	2,1	2	451	0,995576	0,990734
Dentro	21	0,0078					
Entre	120	3,5192	1,22	2	451	0,995576	0,994606
Dentro	121	0,0078					

Analisando a Tabela 4.3, notamos que, a mudança nos graus de liberdade (fixado o número

de classes, e alterando apenas o número de réplicas em cada classe; ou vice-versa) quando $\mathcal{F}$ é fixo, não exerce grande influência na distância entre R e R_I .

Analisando a expressão (4.1), é fácil perceber que as igualdades de $\mathcal{F}$ e número de réplicas implicam na igualdade de R . Observando os resultados da Tabela 4.2 e os da Tabela 4.3, vemos que $\mathcal{F}$ influencia a distância entre R e R_I ; porém, a influência dos graus de liberdade nessa distância é menor.

4.3 Algumas Conseqüências da Não-Confabilidade

Nesta seção, vamos apresentar algumas conseqüências da não-confabilidade quantificando-as como funções de ρ . Através de alguns exemplos, ilustramos como o grau de confiabilidade pode interferir na análise estatística.

4.3.1 Sub-estimação da Correlação

Estamos interessados em estimar a correlação entre duas variáveis T e U , porém em nosso experimento observamos apenas os valores de X e de Y , onde:

$$X = T + \epsilon, \quad e$$

$$Y = U + \eta,$$

ϵ e η são erros aleatórios não-correlacionados entre si, e não-correlacionados com T e U .

Sejam ρ_{XY} e ρ_{TU} os coeficientes de correlação entre X e Y e entre T e U , respectivamente. Com essas considerações, obtemos a seguinte relação:

$$\rho_{XY} = \rho_{TU} \sqrt{R_X R_Y}, \quad (4.3)$$

onde R_X e R_Y são as confiabilidades estimadas de X e de Y , respectivamente.

Como, possivelmente, $\sqrt{R_X R_Y}$ é menor do que 1, temos $\rho_{XY} < \rho_{TU}$, implicando que a correlação entre T e U é sub-estimada.

Notemos que, quanto menor a confiabilidade dos dados, mais afastada estará a correlação observada ρ_{XY} da correlação verdadeira ρ_{TV} . Isso pode nos levar a rejeitar a hipótese nula $H_0 : \rho_{TV} = \rho_0$ (ρ_0 fixado, tal que $\rho_0 \in [0, 1]$), quando na realidade esta é verdadeira.

4.3.2 Problemas na Análise Estatística de Ensaios Clínicos

Desejamos comparar o efeito de 2 tratamentos. O resultado é baseado em uma medida de cada paciente, e de um estudo anterior, sabemos que a confiabilidade estimada das medidas é R . Suponhamos que as variâncias entre os pacientes dos 2 grupos são iguais ($\sigma_{a1}^2 = \sigma_{a2}^2$). Vamos ver qual a influência que a não-confiabilidade pode exercer no estudo.

Tamanho da Amostra

Ao comparar dois tratamentos, o tamanho da amostra necessário para rejeitarmos a hipótese nula $H_0 : \mu_1 - \mu_2 = \delta$, ao nível de significância α , e com poder do teste igual a $1 - \beta$, é dado pela seguinte expressão:

$$N^* = \frac{2\sigma_y^2(Z_{\frac{\alpha}{2}} + Z_{\beta})^2}{\delta^2} \quad (4.4)$$

onde Z_c é tal que $P(Z \leq Z_c) = c$ e $Z \sim Normal(0, 1)$.

Quando não houver erro aleatório nas observações, $\sigma_y^2 = \sigma_a^2$. No entanto, vemos que, se existirem erros aleatórios nas observações, a variância usada para determinar o tamanho da amostra será a soma dos 2 componentes. Podemos escrever esta variância como função de ρ .

De fato, partindo da definição de ρ , obtemos:

$$\sigma_e^2 = \frac{1 - \rho}{\rho} \sigma_a^2$$

Então

$$\sigma_y^2 = \sigma_a^2 + \sigma_e^2 = \frac{1}{\rho} \sigma_a^2 \quad (4.5)$$

e o tamanho da amostra passará a ser:

$$N = \frac{2(\sigma_a^2 + \sigma_e^2)(Z_{\frac{\alpha}{2}} + Z_{\beta})^2}{\delta^2} = \frac{N^*}{\rho} \quad (4.6)$$

No caso em que ρ é desconhecido, usamos a sua estimativa R .

Notemos que, N é frequentemente maior que N^* .

Para ilustrarmos este fato, suponhamos que a mudança da pressão diastólica seja usada para comparar dois tratamentos aplicados em dois grupos independentes de pacientes. Se a hipótese a ser testada é a de que a diferença entre os tratamentos é maior que 5 mmHg, ao nível de significância 5 %, com o poder de teste igual a 80%, e supondo que o desvio padrão da pressão entre os indivíduos seja 8 mmHg obtemos

$$N^* = \frac{2(8)^2(1,96 + 0,842)^2}{5^2} = 40$$

Ou seja, se não existir erro nas observações, necessitamos de 80 pacientes para realizar o estudo (40 pacientes em cada grupo). Porém, se a confiabilidade na medida de pressão for $R = 0,67$, o número de pacientes requisitado em cada grupo é $n = 60$ e o tamanho total da amostra passa a ser 120, isto é, necessitamos de uma amostra 50% maior que a do estudo original. E, indo além, se a confiabilidade for igual a 0,80, a amostra requisitada total será de tamanho 100, ou seja, 25% maior do que a original. Notamos então que, quanto menor a confiabilidade dos dados, maior o tamanho da amostra e, conseqüentemente maior o custo da pesquisa.

Teste de Hipóteses

Na comparação de 2 tratamentos, considerando que os 2 grupos em estudo têm a mesma variância, o teste usual para verificarmos a igualdade das médias é dado pela estatística t^* ,

$$t^* = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_p^2 \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(N_1-1)S_1^2 + (N_2-1)S_2^2}{N_1+N_2-2} \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}} \quad (4.7)$$

onde $\bar{x}_i$ é a média amostral do grupo i ,

S_i^2 é a variância amostral não-viesada do grupo i ,

N_i é o tamanho da amostra do grupo i ($i=1,2$).

S_p^2 é a variância comum dos dois grupos.

Este teste é baseado na suposição de inexistência do erro aleatório. Neste caso, S_p^2 que estima a variância comum dos dois grupos é calculada utilizando S_1^2 e S_2^2 (onde $S_1^2 = \hat{\sigma}_{a1}^2$ e

$S_2^2 = \hat{\sigma}_{a_2}^2$). Caso esta suposição não seja válida, será necessário corrigir (4.7), introduzindo a variação devido ao erro aleatório na estimação da variância comum. Como só temos uma medida por indivíduo, torna-se inviável a estimação de σ_e^2 . Para contornar este problema, podemos utilizar a confiabilidade estimada do estudo anterior. Substituindo os parâmetros de (4.5) por suas respectivas estimativas, temos

$$\widehat{\sigma_{yi}^2} = \frac{1}{R} \widehat{\sigma_{ai}^2} = \frac{1}{R} S_i^2.$$

Assim,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(N_1-1)\frac{1}{R}S_1^2 + (N_2-1)\frac{1}{R}S_2^2}{N_1+N_2-2} \left(\frac{1}{N_1} + \frac{1}{N_2}\right)}} = t^* \sqrt{R} \quad (4.8)$$

Observamos que, para $R < 1$, t é menor do que t^* . Isso significa que, se a correção não for feita, toda vez que utilizamos t^* , corremos o risco de super-estimar o valor de t e, portanto, chegar a conclusões erradas.

É importante ressaltarmos que, esta correção só pode ser aplicada quando os dois grupos têm a mesma variância (isto é, $\sigma_{a_1}^2 = \sigma_{a_2}^2$), caso contrário, a formulação do teste seria outra, e a correção deveria ser feita com a confiabilidade de cada grupo.

Intervalo de Confiança

A construção do intervalo de confiança para a média é baseada na seguinte expressão

$$I.C.: \bar{X} \pm t_n \frac{S}{\sqrt{n}} \quad (4.9)$$

onde $\bar{X}$ é a média amostral,

S é o desvio amostral,

n o tamanho da amostra,

t_n é o valor tabelado da distribuição t-student, com n graus de liberdade.

Quando não existir erro nas medidas, S é uma estimativa de σ_y . Pelo que vimos antes, a correção deve ser feita com a substituição de $\frac{S}{\sqrt{R}}$ no lugar de S em (4.9), (onde R é a confiabilidade estimada pelo estudo anterior). Com isso, levamos em consideração a variação

devido ao erro aleatório que não conseguimos estimar no presente estudo (por só termos uma observação por classe). Assim, o novo intervalo de confiança é determinado por

$$I.C.: \bar{X} \pm t_n \frac{S}{\sqrt{Rn}} \quad (4.10)$$

Notemos que, o intervalo apresentado em (4.10) é, frequentemente, maior do que o intervalo obtido por (4.9). Além disso, quanto maior for R , mais próximo será este intervalo ao intervalo anterior.

Aumento de Viéses para Amostra Não-aleatória

Uma estratégia comumente utilizada para a comparação de dois tratamentos clínicos - ou mesmo em outros tipos de estudos clínicos - é, recrutar para a análise, pacientes que tenham escores acima de um determinado valor limite pré-estabelecido para a variável de interesse. Uma justificativa para o uso deste procedimento é que considerando pacientes com sintomas mais evidentes teremos, possivelmente, mais clareza na detecção do efeito do tratamento.

Sejam M o valor limite adotado, μ a média populacional da variável de interesse e σ_a o desvio padrão dos escores verdadeiros. O ideal seria admitir para a análise somente aqueles pacientes cujos escores verdadeiros T fossem maiores que M . Contudo, devido à impossibilidade de conhecermos T , em situações práticas, um paciente pode ser admitido se o seu escore observado $X = T + \epsilon$ é maior que M . Atuando desta forma, a amostra pode conter pacientes que tecnicamente não deveriam ser incluídos na amostra, pois seus escores verdadeiros T são menores que M , ou seja, são falsos-positivos. Sua presença na amostra se dá pela presença do erro aleatório na medida.

Se a atribuição dos pacientes aos tratamentos for completamente aleatória, nenhum viés é introduzido no estudo pelos pacientes falsos-positivos na comparação dos tratamentos, porém, a amostra total é viesada por conter pacientes falsos-positivos.

Denominemos por P a razão de falsos-positivos, isto é, o número de pacientes falsos-positivos em relação ao número total de pacientes e definimos $D = \frac{(M-\mu)}{\sigma_a}$ como a distância padronizada entre o escore limite e a média populacional. A Tabela 4.4 a seguir mostra a

razão de falsos-positivos em função de D e da confiabilidade estimada R, supondo que D e R têm distribuições normais. Esta tabela foi construída usando-se a Tabela da Normal Bivariada publicada pela National Bureau of Standards, 1959 (Fleiss,1986).

Tabela 4.4: Razão de Falsos-Positivos para a seleção de uma amostra na base de distância padronizada D.

D	Confiabilidade (R)				
	,50	,75	,85	,90	,95
-2,0	,01	,01	,01	0	0
-1,5	,03	0,2	,02	,01	,01
-1,0	,07	,05	,04	,03	,02
-0,5	,14	,09	,07	,06	,04
0	,25	,17	,12	,10	,07
0,5	,40	,26	,20	,16	,11
1,0	,55	,37	,28	,22	,15
1,5	,70	,49	,39	,30	,21
2,0	,82	,61	,47	,38	,26

Notamos que, para R fixo, a razão de falsos-positivos aumenta à medida em que D aumenta. Este fato é intuitivo e pode ser explicado facilmente: D indica a distância entre o escore limite adotado e a média populacional. Quanto maior for D, mais distante estará M de μ , e conseqüentemente, diremos que a maioria dos pacientes apresentam medidas abaixo do limite M fixado, restando poucos para serem incluídos na análise. E, além disso, estes terão chance muito alta de serem falsos-positivos. Contrariamente, se $M < \mu$, D é negativo, e assim, a maioria dos pacientes poderá fazer parte da amostra e com pouca chance de serem falsos-

positivos.

Por outro lado, se fixarmos D a razão de falsos-positivos diminui lentamente quando a confiabilidade aumenta. Por exemplo, suponhamos que seja estabelecido a priori que a proporção de indivíduos falsos-positivos P não possa ultrapassar 16 %. Observamos na Tabela 4.4, que considerando $D=1,0$ apenas com $R=0,95$, esta exigência estará satisfeita. Observamos também que, para D maior do que 0, mesmo com confiabilidade consideravelmente alta, ainda temos uma proporção indesejável de pacientes possíveis falsos-positivos na amostra. Veja, por exemplo, o caso em que $D=1,0$ e $R=0,90$.

4.4 Controle da Confiabilidade pelo Número de Réplicas

Em confiabilidade bem como em toda a teoria estatística, análises baseadas em um conjunto de observações com várias réplicas são sempre mais "confiáveis" do que aquelas feitas considerando-se apenas uma única réplica. Por exemplo, se você pretende estimar a taxa de colesterol de um indivíduo, certamente, você preferirá utilizar a média de várias medidas desta taxa de colesterol à uma única medida, pois, intuitivamente, achamos que a informação trazida pela média é mais "precisa".

Este aumento de confiança pode ser quantificado através do coeficiente de confiabilidade da seguinte maneira.

Seja $\overline{y_m}$, a média das m observações de uma classe. Como a covariância entre as observações de uma mesma classe é σ_a^2 , é fácil de verificar que

$$Var(\overline{y_m}) = \sigma_a^2 + \frac{\sigma_e^2}{m}$$

Com isso e a definição de ρ , obtemos

$$\rho_m = \frac{\sigma_a^2}{Var(\overline{y_m})} = \frac{m\rho}{1 + (m-1)\rho} \quad (4.11)$$

onde ρ_m é a confiabilidade da média de m réplicas para cada indivíduo. Quando ρ é desconhecido substituímos ρ por R - a confiabilidade de uma única medida. Assim, a confiabilidade da média pode ser determinada como uma função da confiabilidade de uma única réplica (Fleiss, 1986).

Utilizando a substituição de ρ por R e observando a expressão (4.11), verificamos ainda que $R_m \geq R$, ou seja, a confiabilidade para a média das réplicas, na pior das hipóteses, é igual a confiabilidade de uma única medida. Claramente, para R diferente de zero, essa igualdade só acontecerá quando tivermos confiabilidade de 100% nas medidas, ou quando $m = 1$.

Suponhamos que o pesquisador estabeleça antes da realização do experimento que a confiabilidade da média das réplicas deva ser R^* . De (4.11) podemos determinar o número de réplicas necessário para que essa meta possa ser atingida. Sendo R a estimativa de ρ , temos

$$m = \frac{R^*(1 - R)}{R(1 - R^*)} \quad (4.12)$$

Por exemplo, se a confiabilidade estimada é $R = 0,75$, e se o pesquisador pré-estabeleceu $R^* = 0,90$, o número de réplicas m necessário para cada indivíduo em um estudo posterior deve ser 3.

4.5 Incorporando o custo

Um dos principais fatores de limitação na realização de um experimento é o custo envolvido. Em alguns experimentos este custo é muito baixo, possibilitando a seleção de uma amostra grande, ou a realização de muitas réplicas para cada elemento da amostra. Em outros, o custo é muito elevado o que limita o experimento à escolha de uma amostra pequena e/ou com poucas réplicas. Estes dois casos extremos não serão considerados nesta seção. Tentaremos exibir aqui uma relação entre o custo do experimento e a confiabilidade, que forneça o número de réplicas necessário para que o experimento tenha custo mínimo, quando uma dada confiabilidade é desejada.

Vamos assumir que o custo total de inclusão de um indivíduo novo na análise seja uma

unidade monetária, e c o custo relativo de obtenção de uma única medida de um dado indivíduo sobre o custo total do experimento. Então, o custo de apenas uma medida para este indivíduo será $1+c$. Consideremos um estudo planejado para a comparação de dois tratamentos. Sejam R a confiabilidade estimada para uma única medida e $N_{(1)}$ o tamanho da amostra para cada grupo fornecido por (4.6) da seção 4.3.2. Segundo Fleiss(1986), se tomarmos m réplicas para calcular a média de cada indivíduo, o tamanho da amostra necessário (denotado por $N_{(m)}$) para cada grupo é:

$$N_{(m)} = \frac{N_{(1)}[1 + (m-1)\rho]}{m}. \quad (4.13)$$

É fácil ver que, como $\rho \leq 1$, $\frac{1+(m-1)\rho}{m} \leq 1$, e portanto $N_{(m)}$ é sempre menor ou igual do que $N_{(1)}$ quando $m \geq 1$. O custo da inclusão de um novo indivíduo aumenta para $1 + mc$. Uma análise descuidada desta observação levará à conclusão de que aumentando o número de réplicas teremos um aumento no custo total do experimento. Olhando mais atentamente percebermos que isso não é necessariamente verdade, como podemos observar a seguir:

O custo total (C_T) para o experimento é

$$C_T(m) = 2 \times N_{(m)}(1 + cm) = 2N_{(1)} \frac{[1 + (m-1)\rho]}{m} (1 + cm)$$

que é proporcional a

$$\frac{[1 + (m-1)\rho](1 + cm)}{m}. \quad (4.14)$$

Derivando (4.14) em relação a m , e igualando a expressão obtida a zero, obtemos

$$m = \sqrt{\frac{1-R}{cR}}. \quad (4.15)$$

Como a segunda derivada de (4.14) é maior do que zero, a expressão dada em (4.15) fornece o ponto que minimiza (4.14). Se m em (4.15) não for inteiro, os dois inteiros mais próximos dele são considerados, e o número ótimo de réplicas por indivíduo será, dentre eles, aquele que fornece o menor valor para a expressão (4.14).

Como m pode ser maior que um, fica claro então que o custo total do experimento pode ser minimizado para valores de m maiores que um, justificando assim o uso de réplicas no

experimento.

Consideremos $R = 0,75$ e $\widehat{\sigma}_a^2 + \widehat{\sigma}_e^2 = 25,77$ e suponhamos que um estudo seja planejado para comparar 2 tratamentos, com poder de teste $1 - \beta = 0,80$, nível de significância $\alpha = 0,05$ e a diferença entre os dois tratamentos $\delta = 2,5$. Com estas considerações e, optando por um experimento com uma réplica, pela expressão (4.6) obtemos:

$$N_{(1)} = 65.$$

Considerando que o custo para uma medida (c) é igual a 0,1, o custo total do experimento será

$$C_T(1) = 2N_{(1)}(1 + c) = 143$$

unidades monetárias.

Consideremos agora um experimento com mais de uma réplica. De (4.15), sabemos que o custo total é minimizado quando

$$m = 1,8.$$

Os valores inteiros de m mais próximos de 1,8 são $m = 1$ e $m = 2$. Sendo assim, substituindo $m = 1$ e $m = 2$ em (4.14) obtemos

$$C_T(1) \propto 1,1$$

$$\text{e } C_T(2) \propto 1,05.$$

Ou seja, o custo é minimizado quando $m = 2$. E, além disso, o número de indivíduos necessário em cada grupo, fornecido pela expressão (4.13) é

$$N_{(2)} = 57.$$

Consequentemente, o custo total do experimento, neste caso, é

$$C_T = 2N_{(2)}(1 + 2c) = 137$$

Verificamos então que, com a introdução de mais réplicas podemos obter uma diminuição tanto no número de indivíduos na amostra, quanto no custo total do experimento. Obviamente, em algumas situações, a utilização de mais do que uma réplica pode conduzir a um custo maior do experimento. De fato, isso ocorre sempre que a seguinte relação entre o custo por medida e a confiabilidade for verificada

$$c > \frac{(1 - R)}{R}$$

No exemplo acima, se o custo por medida for maior ou igual a $1/3$, o experimento com menor custo é aquele com apenas uma réplica.

Podemos observar também que, quando R se aproxima de 1, qualquer que seja o valor de c , a adoção de uma réplica por indivíduo, minimizará sempre o custo total, economizando tempo e dinheiro.

4.6 Estudo de Confiabilidade Inter-juízes

Em seções anteriores vimos como estimar e controlar a confiabilidade de uma medida considerando apenas que as réplicas tenham sido obtidas em condições similares. Por exemplo, consideramos que as amostras de sangue dos pacientes, colhidas dia-a-dia, eram analisadas sempre por um mesmo laboratório, ou por laboratórios com condições similares de análise. Essa seção é dedicada a uma situação um "pouco" mais complexa, onde as réplicas podem ser também explicadas por essas diferenças de condições. Um exemplo disso é supor que uma amostra de sangue de cada paciente é analisada por diversos laboratórios. Assim, as medidas poderiam também diferir devido às diferenças entre a qualidade dos aparelhos usados nos laboratórios. Quando defrontamos com esta situação, além do interesse em avaliar o grau de confiabilidade da medida, podemos também querer verificar a influência, por exemplo, do fator laboratório nesta avaliação.

Esse tipo de estudo tem recebido maior atenção em Psiquiatria (Fleiss, 1986). Comumente, nesta área de estudos, estamos interessados em avaliar o comportamento de pacientes e, para isso, muitas vezes consideramos a opinião de vários psiquiatras - fator que pode influenciar na

avaliação do grau de confiabilidade - cabe lembrarmos que os psiquiatras não necessariamente fazem suas avaliações no mesmo local e à mesma hora. A única exigência é que todos eles tenham acesso às mesmas informações.

Esse fator de influência será denominado "juízes" e os dados obtidos neste tipo de estudo poderão ser organizados na forma da Tabela 4.5.

Tabela 4.5: Avaliação de pacientes por n juízes

pacientes	juízes				
	1	...	j	...	n
1	x_{11}	...	x_{1j}	...	x_{1n}
.	...	...	...	...	...
i	x_{i1}	...	x_{ij}	...	x_{in}
.	...	...	...	...	...
k	x_{k1}	...	x_{kj}	...	x_{kn}

Claramente, a obtenção de alta confiabilidade neste tipo de estudo envolve, não apenas o controle do erro aleatório associado às medidas, mas também aos julgamentos dos juízes. Tanto maior será a confiabilidade, quanto mais concordantes forem tais julgamentos (acreditando-se, obviamente que o erro aleatório seja pequeno).

Segundo Bartko(1976), existem várias maneiras de avaliarmos a confiabilidade de uma observação. A escolha do método deve ser baseada no número de juízes e no tipo de variável em questão. Aqui também, um método simples para avaliação de uma variável contínua é o coeficiente de correlação intraclassa estimado através do modelo de efeitos aleatórios. Além da simplicidade de estimação, esta técnica não exige que todos os pacientes sejam avaliados pelo mesmo número de juízes.

O que difere o modelo a ser estudado nesta seção do modelo considerado na seção 4.1 é a inclusão do fator "juízes" e, conseqüentemente, da interação entre este fator e o fator

”indivíduos”.

Se o grupo de juízes envolvidos no estudo for o mesmo, tanto para o estudo de confiabilidade, quanto para o estudo principal a ser realizado no futuro, o fator ”juízes” é considerado um fator fixo. Caso contrário (isto é, quando os juízes formam uma amostra aleatória da população de juízes), o fator é aleatório.

Apresentamos a seguir a estimação de ρ nas situações citadas acima e também alguma discussão sobre a interação dos fatores considerados no modelo.

4.6.1 Modelo com Dois Fatores Aleatórios

Neste modelo, os juízes envolvidos no estudo constituem uma amostra da população de juízes e então, o modelo que adotamos é:

$$Y_{ij} = \mu + A_i + B_j + (AB)_{ij} + \epsilon_{ij} \quad i = 1, \dots, k; \quad j = 1, \dots, n \quad (4.16)$$

onde

Y_{ij} : a avaliação do j-ésimo juiz sobre o i-ésimo indivíduo;

μ : a média geral populacional;

A_i : o efeito aleatório do i-ésimo indivíduo;

B_j : o efeito aleatório do j-ésimo juiz;

$(AB)_{ij}$: a interação entre o juiz j e o indivíduo i;

ϵ_{ij} : o erro aleatório na ij-ésima observação.

A_i , B_j , $(AB)_{ij}$ e ϵ_{ij} são completamente independentes entre si com distribuições normais, com médias zero, e variâncias σ_a^2 , σ_b^2 , σ_{ab}^2 , e σ_e^2 , respectivamente.

Sendo assim,

$$Var(Y_{ij}) = Var(\mu + A_i + B_j + (AB)_{ij} + \epsilon_{ij}) = \sigma_a^2 + \sigma_b^2 + \sigma_{ab}^2 + \sigma_e^2, \quad (4.17)$$

ou seja, a variância de Y é a soma de 4 componentes de variância.

O coeficiente de correlação intraclasse é então definido da seguinte maneira¹

$$\rho = \frac{\sigma_a^2}{\sigma_y^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2 + \sigma_{ab}^2 + \sigma_e^2}. \quad (4.18)$$

Também aqui ρ tem o papel de mensurar confiabilidade. Ele exerce o papel de um indicador da "gravidade" de variação nos dados. Quando $\rho \rightarrow 0$, isto é, quando a variação entre os indivíduos for muito menor que as variações ocorridas devido ao erro aleatório, ao fator "juízes" e à interação entre os dois fatores envolvidos no modelo, teremos baixa confiabilidade. Esta constatação de baixa confiabilidade sugere uma melhor investigação sobre as fontes de variação citadas acima para melhorar a qualidade dos dados. Por outro lado, se $\rho \rightarrow 1$, a confiabilidade é alta constatando que apenas a variação no fator "indivíduos" é relevante quando comparada às outras fontes de variação.

Para estimar os componentes de variância, construímos a Tabela 4.6,

Tabela 4.6: ANOVA para o modelo com dois fatores

F.V.	G.L.	SQ	QM	Fator Aleatório	Fator Fixo
				E(QM)	E(QM)
indivíduos	$(k - 1)$	SQA	QMA	$\sigma_e^2 + \sigma_{ab}^2 + n\sigma_a^2$	$\sigma_e^2 + n\sigma_a^2$
juízes	$(n - 1)$	SQB	QMB	$\sigma_e^2 + \sigma_{ab}^2 + k\sigma_b^2$	$\sigma_e^2 + \sigma_{ab}^2 + \frac{k}{n-1} \sum B_j^2$
resíduos	$(k - 1)(n - 1)$	SQR	QMR	$\sigma_e^2 + \sigma_{ab}^2$	$\sigma_e^2 + \sigma_{ab}^2$
total	$nk - 1$				

de onde obtemos os seguintes estimadores para os componentes de variância:

$$\widehat{\sigma_{ab}^2} + \widehat{\sigma_e^2} = QMR$$

$$\widehat{\sigma_a^2} = \frac{(QMA - QMR)}{n}$$

¹A definição de ρ tem que estar de acordo com o modelo adotado, porém nem sempre isso é obedecido. Exemplos e discussões sobre este fato podem ser encontrados em Bartko(1966) e Fleiss(1986).

$$\widehat{\sigma_b^2} = \frac{(QMB - QMR)}{k}$$

Como pode ser visto acima, não é necessário fazer nenhuma suposição sobre a ausência ou presença de interação entre os fatores ao estimar os componentes de variância do modelo. Substituindo estes estimadores em (4.18), temos

$$R = \frac{(QMA - QMR)}{QMA + QMR(n-1) + \frac{n}{k}(QMB - QMR)}. \quad (4.19)$$

Usando a aproximação de Satterthwaite(1946), obtemos o seguinte intervalo aproximado para ρ :

$$\rho > \frac{k(QMA - F_{k-1, v^*, \alpha} QMR)}{kQMA + F_{k-1, v^*, \alpha} [nQMB + (N - k - n)QMR]},$$

onde $F_{k-1, v^*, \alpha}$ é o valor crítico da distribuição F, com $\alpha\%$ na cauda superior, $(k-1)$ e v^* são graus de liberdade, com v^* e $\mathcal{F}$ dados pelas seguintes expressões, respectivamente.

$$v^* = \frac{(k-1)(n-1)(nR\mathcal{F} + k[1 + (n-1)R] - nR)^2}{(k-1)n^2R^2\mathcal{F}^2 + (k[1 + (n-1)R] - nR)^2};$$

$$\mathcal{F} = \frac{QMB}{QMR};$$

Se a confiabilidade estimada (R) for baixa, as idéias desenvolvidas na seção 4.4 podem também ser convenientemente aplicadas aqui, tomando-se as devidas cautelas - isto é, lembrando sempre que o estimador de ρ a ser usado é aquele desenvolvido para o modelo com dois fatores exibido pela expressão (4.19).

4.6.2 Modelo Misto - com um fator fixo

Nesta seção, o fator "juízes" é considerado fixo. O grupo de juízes envolvidos no estudo preliminar para a avaliação da confiabilidade é mantido no estudo principal a ser realizado no futuro.

O modelo adequado para essa situação tem uma estrutura análoga ao modelo apresentado em (4.16) da seção anterior. A diferença é que, agora, B_j é o efeito *fixo* do j -ésimo juiz, com a restrição de que $\sum_j B_j = 0$. Este modelo é conhecido na literatura como *Modelo Misto*.

Neste caso, podemos verificar que a variância de Y_{ij} é soma de 3 componentes de variância, isto é,

$$Var(Y_{ij}) = \sigma_a^2 + \sigma_{ab}^2 + \sigma_e^2, \quad (4.20)$$

Sendo assim, o coeficiente de correlação intraclasse é

$$\rho = \frac{\sigma_a^2}{\sigma_y^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{ab}^2 + \sigma_e^2}. \quad (4.21)$$

A tabela de ANOVA para o Modelo Misto tem a mesma estrutura que a apresentada para o modelo com efeitos aleatórios. A única diferença entre eles está na distribuição dos quadrados médios, a qual ocorre unicamente devido à restrição sobre o efeito fixo (isto é, $\sum_j B_j = 0$). Por este motivo, estes resultados estão conjuntamente apresentados na Tabela 4.6, da seção anterior, de onde podemos obter os seguintes estimadores para os componentes de variância:

$$\widehat{\sigma_e^2} + \widehat{\sigma_{ab}^2} = QMR$$

$$\widehat{\sigma_a^2} = \frac{QMA - QMR + \widehat{\sigma_{ab}^2}}{n}$$

e assim, substituindo estes resultados em (4.21) temos o seguinte estimador para ρ :

$$R = \frac{QMA - QMR + \widehat{\sigma_{ab}^2}}{QMA + QMR(n-1) + \widehat{\sigma_{ab}^2}} \quad (4.22)$$

O cálculo de R fica, portanto, condicionado ao conhecimento do valor da variância da interação, ou de algum conhecimento a priori que nos permita desconsiderar a presença da interação entre os dois fatores. Em geral, para o particular problema onde cada juiz faz apenas uma avaliação de cada indivíduo, é relativamente difícil constatar a significância do efeito da interação. Certamente nenhum teste clássico pode ser aplicado nesta situação. Embora isso ocorra, Bartko(1966) sugeriu um procedimento para construir um intervalo para R , que vamos apresentar a seguir.

Consideremos

$$f(x) = \frac{QMA - QMR + x}{QMA + QMR(n-1) + x}$$

Como $\frac{df(x)}{dx} \geq 0$ para todo x , temos que $f(\cdot)$ é uma função crescente. Então para $0 \leq \widehat{\sigma_{ab}^2} \leq QMR$ (lembrando que, $\widehat{\sigma_{ab}^2} + \widehat{\sigma_e^2} = QMR$), teremos $f(0) \leq f(\widehat{\sigma_{ab}^2}) \leq f(QMR)$, ou seja,

$$R_{INF} = \frac{QMA - QMR}{QMA + QMR(n-1)} < R < \frac{QMA}{QMA + nQMR} = R_{SUP}. \quad (4.23)$$

O intervalo assim construído para R tem R_{INF} como limite inferior e R_{SUP} como limite superior.

Como vimos em seções anteriores, vários fatores podem provocar uma baixa confiabilidade nos dados. Já foram citados, a semelhança entre os indivíduos - que pode levar a uma maior sensibilidade do estimador de ρ - e o efeito de um erro aleatório grande - o qual pode ser diminuído pelo aumento do número de réplicas.

Nessa seção, além destes, existe outro fator - a discordância entre os juízes. Quando consideramos o *Modelo Misto*, a influência deste fator não pode ser detectada por ρ (ver expressão (4.21) nesta subseção). Fleiss(1986) apresentou várias sugestões para contornar este problema.

4.7 Aplicação

Nesta seção, vamos apresentar uma aplicação dos conceitos de confiabilidade, vistos anteriormente, em um estudo sobre a mensuração do nível sérico de colesterol (concentração de colesterol no sangue) do ser humano.

As doenças cardio-vasculares são, ainda hoje, as maiores causadoras de morte em todo o mundo. Um dos importantes fatores de risco para estas doenças é a *hipercolesterolemia*. Um indivíduo é considerado como hipercolesterolêmico, se o seu nível sérico de colesterol total

excede o valor de 220 mg/dL. Na presença de múltiplos fatores de risco, a intervenção para abaixar o nível de colesterol é sempre recomendada; neste caso, o nível de colesterol é visto como um indicador de tratamento. Porém, o nível sérico de colesterol não é um valor constante: além da variação analítica (erros laboratoriais, procedimentos de coleta - incluindo postura do paciente na hora de coleta, etc.), existe também a variação biológica (variação natural do indivíduo, que pode variar devido ao dia, estação ou alimentação, etc.). Segundo Bookstein et al (1990), o coeficiente de variação desta última pode chegar a 12% para um mesmo indivíduo.

Nos últimos anos, a literatura médica vem discutindo a credibilidade de se avaliar o nível de colesterol baseando no resultado de um único exame. Devido às variações citadas acima, recomenda-se pedir mais de um exame em caso de dúvida. Então, conhecer a confiabilidade do resultado do exame é uma consideração importante para o clínico, sendo que, a falta deste conhecimento pode levar a um diagnóstico errado, causando consequentemente, desgastes desnecessários (psicológicos e financeiros) ao paciente.

Com o objetivo de conhecer a acurácia da análise do nível sérico de colesterol do laboratório e a variação natural de indivíduos aparentemente saudáveis, pesquisadores do Hospital Universitário (HU) da Universidade de São Paulo iniciaram um estudo. Uma amostra de 32 voluntários foi coletada. Os voluntários são funcionários do HU. Durante um mês, às terças-feira, os voluntários ficaram em jejum de 12 horas para fazer a coleta de sangue. Depois, cada amostra de sangue coletada foi repartida em 3 sub-amostras, e congeladas até serem analisadas. As análises foram realizadas no laboratório do hospital, onde foram processados por técnicos diferentes com aparelhos automáticos.

Para a análise estatística, consideramos o modelo com 2 fatores aleatórios hierárquicos: indivíduo e dia, e os resultados do exame das três sub-amostras foram considerados como réplicas em cada dia. Apresentamos esta estrutura na Figura 4.1.

Escrevemos então:

$$Y_{ijk} = \mu + A_i + B_{j(i)} + \epsilon_{ijk}$$

Figura 4.1: Estrutura dos dados no estudo de nível sérico de colesterol

onde:

Y_{ijk} : k-ésima observação do indivíduo i no dia j;

μ : média geral do nível sérico de colesterol;

A_i : efeito aleatório do indivíduo i, $A_i \sim Normal(0, \sigma_a^2)$;

$B_{j(i)}$: efeito aleatório do dia j no indivíduo i, $B_{j(i)} \sim Normal(0, \sigma_b^2)$;

ϵ_{ijk} : erro aleatório da k-ésima observação do indivíduo i no dia j, $\epsilon_{ijk} \sim Normal(0, \sigma_e^2)$;

A_i , $B_{j(i)}$ e ϵ_{ijk} são completamente independentes entre si.

Como

$$Var(Y_{ijk}) = \sigma_a^2 + \sigma_b^2 + \sigma_e^2$$

o coeficiente de correlação intraclasse tem a seguinte forma:

$$\rho = \frac{\sigma_a^2}{\sigma_y^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2 + \sigma_e^2} \quad (4.24)$$

Utilizando o pacote computacional SAS, obtemos as seguintes estimativas dos componentes de variância:

$$\widehat{\sigma_a^2} = 666,73; \quad \widehat{\sigma_b^2} = 160,25 \quad e \quad \widehat{\sigma_e^2} = 0,85$$

Substituindo as estimativas dos componentes em (4.22), obtemos a estimativa de ρ ,

$$R = 0,81$$

que é interpretada como confiabilidade de uma medida única de colesterol.

Notamos que, $\widehat{\sigma_e^2}$ é muito pequeno em relação aos outros componentes de variância do modelo, significando a alta precisão do laboratório. Portanto, podemos trabalhar só com uma réplica, desconsiderando a existência das outras réplicas no mesmo dia. Como a variação entre as réplicas é pequena, ela não deve interferir no resultado. Substituímos, então, as três réplicas do mesmo dia pelo valor médio, trabalhando com apenas uma medida em cada dia². Assim, para cada indivíduo, temos 4 medidas independentes.

²Esse procedimento foi tomado para aproveitar ao máximo os dados coletados, embora não seja a única forma que podemos adotar. Por exemplo, podemos escolher qualquer réplica dentre as três do mesmo dia.

Consideramos agora apenas o fator Indivíduo. O modelo é reduzido ao de um fator.

Processando os dados, obtemos

$$\widehat{\sigma}_a^2 = 666,73; \quad \widehat{\sigma}_e^2 = 160,53$$

Substituindo as estimativas dos componentes de variância em (4.1), temos

$$R = 0,81$$

Podemos aumentar a confiabilidade das medidas através da utilização da média, como vimos na seção 4.4. Usando (4.11), construímos a seguinte tabela:

Tabela 4.7: Confiabilidades estimadas da média

m=	R_m
1	0,805
2	0,892
3	0,925
4	0,943
5	0,953
6	0,961

onde m é o número de dias necessários para calcular a média, R_m é a confiabilidade estimada da média de m dias. Então, se avaliarmos o paciente em 4 dias diferentes e calcularmos a média, a confiabilidade estimada dessa medida é 0,943.

Vimos na seção 4.1, que o valor observado Y é a soma de valor verdadeiro T com erro aleatório, isto é

$$Y = T + \epsilon$$

Como $\widehat{\sigma}_e^2=160,53$, o valor verdadeiro da medida estará entre $Y - \widehat{\sigma}_e$ e $Y + \widehat{\sigma}_e$, isto é,

$$T = Y \pm 12,67$$

Capítulo 5

Considerações Finais

Nesta dissertação apresentamos e estendemos alguns resultados de Scheffé(1959), Donner(1986) e Fleiss(1986). Ainda que organizada com o intuito de propiciar um acompanhamento linear e lógico do leitor, cada capítulo dessa dissertação contém informações que podem ser utilizadas com uma certa independência de consulta.

No desenvolvimento deste trabalho, apresentamos vários resultados. Neste capítulo, vamos apresentar apenas alguns comentários gerais.

A adoção do planejamento ótimo, nos traz um grande ganho em termos de precisão, porém a sua determinação depende de uma função dos componentes de variância (θ ou ρ). Caso esta função seja desconhecida, uma pesquisa piloto é útil para resolver o problema. Podemos utilizar $N'/4$ classes com 4 observações em cada classe, onde N' é um número natural múltiplo de 4, fixado de antemão, e que representa o número total de unidades experimentais (Este planejamento fornece uma precisão quase máxima na estimação de ρ , qualquer que seja N' como verificado no capítulo 3). Com este estudo inicial, podemos estimar θ (ou ρ), e então, com essa nova informação, podemos usar o programa PO.COM ou os resultados obtidos nos capítulos 2 e 3 para determinarmos um planejamento ótimo.

No estudo de confiabilidade, o planejamento sugerido acima pode ser também utilizado para avaliar a confiabilidade das medidas.

Quando o interesse é fazer um planejamento para a estimação simultânea dos 3 parâmetros de interesse, esta dissertação não fornece diretamente a solução do problema. No entanto, fornece alguns elementos básicos que permitem adequar o melhor planejamento para a situação. Uma sugestão para este problema é, obter o planejamento ótimo para cada parâmetro de interesse - obtemos assim 3 candidatos ao planejamento ótimo para a estimação simultânea dos 3 parâmetros. Cada planejamento é utilizado para averiguar a variância dos estimadores dos outros parâmetros - (por exemplo, o planejamento ótimo na estimação de θ é investigado utilizando os estimadores de σ_a^2 e ρ). O planejamento que fornece as menores variâncias dos estimadores dos 3 parâmetros deve então ser adotado.

Com relação a aplicação prática desenvolvida no capítulo 4, verificamos que, foram utilizados no estudo 4 repetições por indivíduo para estimar o coeficiente de confiabilidade, o que está exatamente em concordância com o planejamento quase-ótimo. Como o valor encontrado de ρ é maior do que 0,2, concluímos que ρ foi estimado com boa eficiência.

Finalmente, devemos ressaltar que, embora não sendo imediato, é possível estender os resultados deste trabalho para modelos mais complexos, com dados não-balanceados, envolvendo mais fatores, o que permite ampliar a aplicabilidade dos resultados a várias áreas de pesquisa.

Apêndice I - As Diferenças entre o Modelo de Efeitos Aleatórios e o Modelo de Efeitos Fixos

A equação matemática que representa os dois modelos com um fator e dados balanceados é :

$$Y_{ij} = \mu + A_i + \epsilon_{ij} \quad i = 1, \dots, k \quad j = 1, \dots, n$$

Resumimos as informações básicas sobre os modelos de efeitos aleatórios e fixos (ambos com um fator) na seguinte tabela:

Tipo de Fator	Suposições	Hipóteses	Estatística
Fixo	$\epsilon_{ij} \sim Normal(0, \sigma_e^2), indep.$	$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ $\mu_i = \mu + A_i$	$\mathcal{F} = \frac{QMA}{QMR}$
Aleatório	$\epsilon_{ij} \sim Normal(0, \sigma_e^2), indep.$ $A_i \sim Normal(0, \sigma_a^2), indep.$ ϵ_{ij} e A_i indep.	$H_0 : \sigma_a^2 = 0$	

(continue...)

(continuação)

Tipo de fator	Distribuição de $\mathcal{F}$ sob H_0	Distribuição de $\mathcal{F}$ sob H_A
Fixo	$\frac{QMA}{QMR} \sim F_{k-1, N-k}(0)$ Distribuição F de Fisher-Snedecor central com k-1 e N-k graus de liberdade	$\frac{QMA}{QMR} \sim F_{k-1, N-k}(\phi)$ onde $\phi = \frac{1}{\sigma_e^2} \sqrt{\frac{n}{k}} \sum_{i=1}^k (\mu_i - \mu)^2$ Distribuição F não central
Aleatório		$\frac{QMA}{QMR} \sim (1 + n \frac{\sigma_a^2}{\sigma_e^2}) F_{k-1, N-k}(0)$ Proporcional à distribuição F central, com k-1, e N-k graus de liberdade

Apêndice II - Produto de Kronecker - Definição e propriedades

O Produto de Kronecker ($\otimes$) é um produto especial de matrizes. Vimos no Capítulo 1 que, com a utilização deste produto, as notações de matrizes e as operações entre elas são simplificadas. Aqui, vamos apresentar a definição e as propriedades deste produto.

Segundo Rao(1973), o Produto de Kronecker é definido da seguinte forma:

Definição: Sejam $A=(a_{ij})$ e $B=(b_{ij})$ duas matrizes de dimensões $m \times n$ e $p \times q$, respectivamente; onde a_{ij} e b_{ij} são os termos gerais das matrizes A e B. Então, o Produto de Kronecker entre as matrizes A e B,

$$A \otimes B = (a_{ij}B)$$

é uma matriz $mp \times nq$, que possui como partições - as matrizes $(a_{ij}B)$ - em cada posição ij , onde $i=1,...,m$ e $j=1,...,n$.

Consequentemente, temos as seguintes propriedades:

a) $0 \otimes A = A \otimes 0 = 0$

b) $(A_1 + A_2) \otimes B = (A_1 \otimes B) + (A_2 \otimes B)$

c) $A \otimes (B_1 + B_2) = (A \otimes B_1) + (A \otimes B_2)$

d) $aA \otimes bB = abA \otimes B$, a e b são escalares.

e) $A_1 A_2 \otimes B_1 B_2 = (A_1 \otimes B_1)(A_2 \otimes B_2)$

f) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$, se as matrizes inversas de A e B existem.

g) $(A \otimes B)^- = A^- \otimes B^-$, onde A^- e B^- são matrizes inversas generalizadas.

h) $(A \otimes B)' = A' \otimes B'$, A' e B' são matrizes transpostas de A e B .

i) $(A \otimes B)(A^{-1} \otimes B^{-1}) = I$, I é a matriz identidade.

Apêndice III - O Programa PO.COM

- as especificações e listagem

Em estudos de planejamento com alocação ótima, a variância do estimador do parâmetro de interesse é um critério para a seleção do planejamento ótimo. O valor desta variância depende não somente do método de estimação, mas também do valor de alguma função dos componentes de variância, por exemplo, θ ou ρ . Conhecer essa função é fundamental para a determinação do planejamento ótimo, mas nem sempre é possível conseguir estimativas confiáveis para essas funções. Para contornar este problema, elaboramos o programa PO.COM, que tenta fornecer todos os possíveis planejamentos ótimos variando ρ em todo o seu campo de variação¹ (isto é, o intervalo $[0,1]$).

Para utilizar este programa no estudo de planejamento ótimo de um experimento é necessária a definição dos seguintes elementos:

- a) parâmetro de interesse a ser estimado;
- b) método de estimação a ser utilizado;
- c) o tamanho da amostra (N);

Depois de receber essas informações básicas, o programa calcula k_{otimo} para cada valor de ρ . Se k_{otimo} obtido não for inteiro, ele procura os dois divisores de N mais próximos do k_{otimo} , compara as variâncias do estimador na adoção destes dois divisores, e então escolhe aquele que

¹ k_{otimo} pode ser escrito em função de θ ou ρ , porém, todos os cálculos deste programa foram baseados em ρ , já que ele tem um campo de variação limitado, e permite, com isso, estudar todos os possíveis planejamentos. Os valores de θ apresentados na coluna THETA da saída do programa são calculados pela expressão $\theta = \frac{\rho}{1-\rho}$.

produz a menor variância e imprime em KSUGERIDO. O processo de cálculo de k_{otimo} começa com ρ igual a zero, e incrementa 0,01 para iniciar o cálculo do próximo k_{otimo} . O processo se repete, até que KSUGERIDO seja $N/2$.

O arquivo de saída contém todos os possíveis planejamentos e as variâncias correspondentes do estimador. Um planejamento é fornecido quando ele é adotado pela primeira vez. Exemplos de interpretação dos resultados podem ser encontrados na seção 2.3.

O programa foi elaborado em linguagem PASCAL (versão 3.1). Devido às limitações da linguagem, o tamanho máximo da amostra que o programa aceita é 32.000. O número de divisores está pré-definido em 60, mas pode ser aumentada.

A listagem deste programa será apresentada a seguir.

program po;

{Este program fornece os planejamentos ótimos para um determinado tamanho da amostra em todo campo de variação de ρ . Os valores de ρ estudados variam entre 0.01 e 1, com incremento de 0.01. O arquivo de saída fornece os valores de K_{otimo} (calculadas pelas fórmulas apresentada nas seções 2.1.2, 2.2.2, 3.1 e 3.2) e o divisor de N mais próximo deste k_{otimo} que produz a menor variância. As variâncias destes valores de k e o valor de θ correspondente a esse valor de ρ são apresentados também. A parte teorica se encontra nos capítulo 2 e 3 da dissertação}

```
const n1=60;
type SEQ=array[1..n1] of integer;
var
D:integer;           {número de divisores}
RESTO:integer;       {verifica se é divisor}
PLAN:integer;        {planejamento otimo}
DIVMIN ,DIVMAX:integer; {divisores mínimo e máximo}
OPCAO,I:integer;
TELA:integer;
N:real;              {tamanho da amostra}
VAROT,VARDMIN,VARDMAX:real; {variâncias}
DIF:real;            {diferença das variâncias}
RHO,THETA,KOTIMO:real;
DIVISOR:SEQ;
INICIO:boolean;      {verifica se está iniciando um novo tamanho de amostra}
FIM:boolean;         {indica o final do programa}
ATENCAO:boolean;     {o planejamento adotado pode ser o último divisor}
NOMEARQ:string[20];
RESPOSTA:char;
SAIDA:text;
{Este procedimento alerta o usuário no caso de erro de entrada das informações básicas}
procedure CAMPAINHA;
begin
  sound(200);
  delay(100);
  nosound;
end;
{Esta função verifica a entrada das informações}
function CORRETO:boolean;
var
AUX:boolean;
begin
  readln(RESPOSTA);
  AUX:=false;
  case tela of
    1: if RESPOSTA in ['s','S','n','N'] then AUX:=true
       else CAMPAINHA;
    2: if RESPOSTA in ['1','2','3','4'] then AUX:=true
       else CAMPAINHA;
    3: if RESPOSTA in ['1','2'] then AUX:=true
       else CAMPAINHA;
  end;
  CORRETO:=AUX;
end;
{Esta funcao controla o tempo de execução do program no caso de erro.
```



```

O programa será encerrado quando  $\rho$  atinge seu valor máximo 1.}
function ERROU:boolean;
var aux:boolean;
begin
  AUX:=false;
  if (rho=1) then AUX:=true;
  ERROU:=AUX;
end;
{Esta função calcula o valor de KOTIMO}
function KCALC:real;
var AUX:real;
begin
  if OPCAO=1 then
    AUX:=  $N*(2-2*\rho+N*\rho)/(1+N-\rho)$ 
  else if OPCAO=2 then
    AUX:= $(\sqrt{N}*\rho-5*N*\rho+2*N-8*\rho)/(N*\rho+N-3+3*\rho)$ 
  else if OPCAO=31 then AUX:= $N*((N-1)*\rho+2)/((N+1)+(N-1)*\rho)$ 
  else
    if OPCAO=32 then AUX:= $N*\rho/(1+\rho)$ ;

  KCALC:=AUX;
end;

{Esta função calcula a variância do planejamento com k classes}
function VARIAN(K:real):real;
var AUX:real;
begin
  if OPCAO=1 then
    AUX:= $2*(\sqrt{\rho}/(K-1)/\sqrt{1-\rho}+2*\rho*K/N/(K-1)/(1-\rho)+(N-1)*\sqrt{K}/\sqrt{N}/(k-1)/(N-K))$ 
  else if OPCAO=2 then
    AUX:= $2*(N-3)*\sqrt{N*\rho+K*(1-\rho)} / \sqrt{N}/(K-1)/(N-K-4)/\sqrt{1-\rho}$ 
  else if OPCAO=31 then AUX:= $2*\sqrt{1-\rho}*\sqrt{K+(N-K)*\rho}/N/(N-K)/(K-1)$ 
  else if (OPCAO=32) and (K<>0) then
    AUX:= $2*\sqrt{1-\rho}*\sqrt{K+(N-K)*\rho}/N/(N-K)/K$ 
  else if K=0 then AUX:=0;

  VARIAN:=AUX;
end;

{Esta função verifica se estamos usando o último divisor}
function ULTIMO:boolean;
var aux:boolean;
begin
  AUX:=false;
  if I=D then AUX:=true;
  ULTIMO:=AUX;
end;

{Este procedimento imprime os resultados no arquivo de saída pré-definido}
procedure IMPRIME(DIVISOR:integer; VARIANCIA:real);
begin
  THETA:= $\rho/(1-\rho)$ ;
  write(SAIDA,' ', THETA:3:2, ' ', RHO:3:2, ' ', KOTIMO:5:2);
  write(SAIDA, ' ', DIVISOR, ' ', VAROT:8:4);
  writeln(SAIDA, ' ', VARIANCIA:8:4);
end;

```

```

{Este procedimento determina os divisores do valor N}
procedure DIVISORES;
var
I,RESTO,AUX:integer;
begin
  AUX:=2;
  I:=0;
  while AUX<=N/2 do
  begin
    RESTO:= trunc(N) mod AUX;
    if RESTO= 0 then begin I:=I+1;
                        DIVISOR[I]:=AUX;
                      end;
    AUX:=AUX+1;
  end;
  D:=I;
end;
{Este procedimento apresenta a primeira tela, faz uma pequena introdução sobre
a utilização deste programa}
procedure INICIALIZA;
begin
  TELA:=1;
  ClrScr;
  writeln('*****');
  writeln('* Este program fornece os planejamentos ótimos que *');
  writeln('* produzem as menores variâncias na estimação de de- *');
  writeln('* terminado parâmetro para um dado tamanho de amos- *');
  writeln('* tra N. *');
  writeln('*****');
  write('Continue? (S/N) ');
  while not CORRETO do write('Continue? (S/N) ');
  if (RESPOSTA='N') or (RESPOSTA='n') then FIM:=true;
end;

{Este procedimento apresenta a tela 2 que permite fazer as escolhas}
procedure ESCOLHA;
var AUX:integer;
begin
  ClrScr;
  TELA:=2;
  writeln('Qual o parâmetro a ser estimado?');
  writeln;
  writeln('1) variância dos efeitos das classes (sigma_a^2)');
  writeln;
  writeln('2) theta');
  writeln;
  writeln('3) rho ');
  writeln;
  writeln('4) SAIR do programa ');
  writeln;
  write('OPÇÃO:');
  while not CORRETO do begin
    writeln('ERRO!! DIGITE NOVAMENTE. ');
    write('OPÇÃO:');
  end;
end;

```

```

writeln;
val(RESPOSTA,OPCAO,AUX);
if OPCA0=3 then begin
    TELA:=3;
    writeln('Qual o método de estimação?');
    writeln('    1) Análise de Variância ');
    writeln('    2) Máxima Verossimilhança ');
    write('OPÇÃO: ');
    while not CORRETO do begin
        writeln('ERRO!! DIGITE 1 OU 2. ');
        write('OPÇÃO: ');
    end;
    if RESPOSTA='1' then OPCA0:=31
    else OPCA0:=32;
end
end;

{ Leitura das informações básicas}
procedure ENTRADA;
begin
    ClrScr;
    writeln('*****');
    writeln('*  Entrada de Informações Básicas  *');
    writeln('*****');

    writeln('Qual o nome do arquivo de saída? Exemplo: teste.res');
    readln(NOMEARQ);
    assign(SAIDA, NOMEARQ);
    rewrite(SAIDA);

    write(SAIDA, 'O parâmetro a ser estimado é ');
    if (OPCA0=1) then
        begin
            writeln(SAIDA, 'a variância dos efeitos das classes');
            writeln(SAIDA);
            writeln(SAIDA, '*** Observação: as variâncias apresentadas são proporcionais a variância');
            writeln(SAIDA, 'apresentada pela expressão (2.4). Para obter os valores exatos das variâncias,');
            writeln(SAIDA, 'é necessário multiplicar os valores apresentados pela variância dos erros alea-');
            writeln(SAIDA, 'tórios elevada a 4 ***');
            writeln(SAIDA);
        end
    else if (OPCA0=2) then writeln(SAIDA, 'THETA')
    else if (OPCA0=31) then begin
        writeln(SAIDA, 'RHO');
        writeln(SAIDA, 'O MÉTODO DE ESTIMAÇÃO É DE ANÁLISE DE
VARIÂNCIA');
    end
    else if (OPCA0=32) then begin
        writeln(SAIDA, 'RHO');
        writeln(SAIDA, 'O MÉTODO DE ESTIMAÇÃO É DE MÁXIMA
VEROSSIMILHANÇA');
    end;

    if INICIO then
        begin

```

```

write('Qual o tamanho da amostra (N)?');
readln(N);
DIVISORES;
writeln;
end;
if D <> 0 then { N não é número primo}
begin
  writeln(SAIDA, 'O tamanho da amostra é: ', N:6:0);
  write(SAIDA, 'Os divisores de N são: ');
  for I:=1 to D do write(SAIDA,DIVISOR[I], ' ');
  writeln(SAIDA);
  writeln(SAIDA);
  write(SAIDA, 'THETA ', 'RHO ', 'KOTIMO ', 'K SUGERIDO ');
  writeln(SAIDA, 'VARIANCIA DO KOTIMO ', 'VARIANCIA DO K ');
end
else begin
  writeln('N é número primo, não existe divisor! ');
  FIM:=true;
end;
end;
{ Início da busca de planejamento ótimo }
procedure Calcule;
begin
  ClrScr;
  writeln('*** Aguarde, processando... ***');
  ATENCAO:=false;
  RHO:=0;
  PLAN:=0;
  DIVMIN:=DIVISOR[1];
  DIVMAX:=DIVISOR[2];
  I:=2;
  repeat begin
    while not ULTIMO or not ATENCAO do
      begin
        KOTIMO:=KCALC;
        VAROT:=VARIAN(KOTIMO);
        if (KOTIMO>DIVMIN) and (KOTIMO < DIVMAX) then
          begin
            VARDMIN:=VARIAN(DIVMIN);
            VARDMAX:=VARIAN(DIVMAX);
            DIF:=VARDMIN-VARDMAX;
            if (DIF >0) and (PLAN<>DIVMAX) then begin
              PLAN:=DIVMAX;
              IMPRIME(DIVMAX, VARDMAX);
              ATENCAO:=true;
            end
          else if (DIF<0) and (PLAN<>DIVMIN) then begin
              PLAN:=DIVMIN;
              IMPRIME(DIVMIN, VARDMIN);
              ATENCAO:=false;
            end;
          end
        else if (KOTIMO> DIVMAX) and (KOTIMO< N/2) then
          begin
            DIVMIN:=DIVMAX;
            I:=I+1;

```

```

        DIVMAX:=DIVISOR[I];
        ATENCAO:=false;
        RHO:=RHO-0.01;
    end
    else if KOTIMO=DIVMIN then begin
        PLAN:=DIVMIN;
        VARDMIN:=VAROT;
        IMPRIME(DIVMIN,VARDMIN);
        ATENCAO:=false;
    end
    else if (KOTIMO=DIVMAX) or (KOTIMO > N/2) then
        begin
            PLAN:=DIVMAX;
            VARDMIN:=VAROT;
            IMPRIME(DIVMAX,VARDMAX);
            ATENCAO:=true;
        end;

    RHO:=RHO+0.01;
end;
end
until ERROU or ULTIMO;
write('**** Fim da execução. Os resultados serão encontrados no arquivo ');
writeln(NOMEARQ);
writeln;
end;

```

```

{ ***** Program Principal ***** }

```

```

begin
    FIM:=false;
    INICIO:=true;
    INICIALIZA;
    while not FIM do
        begin
            ESCOLHA;
            if (OPCAO <> 4) then
                begin
                    if INICIO then begin
                        ENTRADA;
                        INICIO:=false;
                    end
                end
            else begin
                TELA:=1;
                writeln('O tamanho da amostra anterior é ',N:6:0);
                write('Usa o mesmo tamanho da amostra? (S/N)');
                while not CORRETO do write('Erro!! Responda S ou N: ');
                writeln;
                if (RESPOSTA='n') or (RESPOSTA='N') then INICIO:=true;
                ENTRADA;
                INICIO:=false;
            end;
        end
        if not FIM then CALCULE;
        TELA:=1;
        write('Deseja finalizar? (S/N)');
        while not CORRETO do write('Erro!! Responda S ou N:');
    end

```



```
writeln;  
if (RESPOSTA='s') or (RESPOSTA='S') then FIM:=true  
else FIM:=false;  
end  
else FIM:=true;  
close(SAIDA);  
end;  
end.
```

Bibliografia

- [1] Airy, B.B.(1861) **On the Algebraical and Numerical Theory of Errors of Observations and the Combinations of Observations**. London: MacMillan.
- [2] Anderson, R.L. & Bancroft, T.A.(1952) **Statistical Theory in Research**. New York: McGraw-Hill.
- [3] Bartko, J.J.(1966) The intraclass correlation coefficient as a measure of reliability. **Psychological Reports** **19**, 3-11.
- [4] Bartko, J.J. & Carpenter, W.T.(1976) On the methods and theory of reliability. **The Journal of Nervous and Mental Disease** **163**, 5, 307-317.
- [5] Bezerra, R.C.F.(1976) **Componentes de Variância - O Problema de Estimativas Negativas**. São Paulo: IME-USP.
- [6] Bookstein, L., Gidding, S. S., Donovan, M. and Smith, F.A.(1990) Day-to-day variability of serum cholesterol, triglyceride, and high-density lipoprotein cholesterol levels. **Arch. Intern. Med.** **150**, August, 1653-1657.
- [7] Box, G.E.P., Hunter, W.G. & Hunter, J.(1978) **Statistics for Experiments. An Introduction to Design, Data Analysis and Model Building** New York: John Wiley & Sons.
- [8] Brownlee, K.A. (1965) **Statistical Theory and Methodology: In Science and Engineering**. New York: John Wiley & Sons.

- [9] Bulmer, M.G. (1957) Approximate confidence limits for components of variance. **Biometrika** **44**, 159-167.
- [10] Chauvenet, W.(1863) **A Manual of Spherical and Practical Astronomy, 2: Theory and Use of Astronomical Instruments**. Philadelphia: Lippincott.
- [11] Cochran, W.G.(1939) The use of analysis of variance in enumeration by sampling. **J. Am. Statist. Assoc.** **34**, 492-510.
- [12] Crump, S.L.(1947) **The Estimation of Components of Variance in Multiple Classification**. Ph. D. Thesis, Iowa State University, Ames, Iowa.
- [13] Daniels, H.E.(1939) The estimation of components of variance. **J. R. Statist. Soc. Suppl.** **6**, 186-197.
- [14] Donner, A.(1986) A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model. **International Statistical Review** **54**, 1, 67-82.
- [15] Donner, A. & Koval, J.J.(1980) The large sample variance of an intraclass correlation **Biometrika** **67**, 719-722.
- [16] Dunn, O.J.(1981) One-way analysis of variance. In: **Encyclopedia of Statistical Sciences volume 9**. New York: John Wiley & Sons.
- [17] Eisenhart, C.(1947) The assumptions underlying the analysis of variance. **Biometrics** **3**, 1, 1-21.
- [18] Fisher, R.A.(1918) The correlation between relatives on the supposition of mendelian inheritance. **Trans. R. Soc., Edinburgh** **52**, 399-433.
- [19] Fisher, R.A.(1925) **Statistical Methods for Research Workers**. London: Oliver and Boyd.

- [20] Fleiss, J.L.(1986) **The Design and Analysis of Clinical Experiments**. New York:Jonh Wiley & Sons.
- [21] Ganguli, M.(1941) A note on nested sampling. **Sankhyã** 5, 449-452.
- [22] Graybill, F.A.(1961) **An Introduction to Linear Statistical Models. Volume I**. New York: McGraw-Hill.
- [23] Graybill, F.A. & Hultquist, R.A.(1961) Theorems concerning Eisenhart's Model II. **Ann. Math. Statist.** 32, 261-269.
- [24] Graybill, F.A. & Wortham, A.W.(1956) A note on uniformly best unbiased estimators for variance components. **J. Am. Statist. Assoc.** 51, 266-268.
- [25] Hammersley, J.M.(1949) The unbiased estimate and standard error of the interclass variance. **Metron** 15, 189-205.
- [26] Hartley, H.O. & Rao, J.N.K.(1967) Maximum likelihood estimation for the mixed analysis of variance model. **Biometrika** 54, 93-108.
- [27] Hultquist, R.(1981) Variance Components. In: **Encyclopedia of Statistical Sciences volume 9**. New York: John Wiley & Sons.
- [28] Jackson, R.W.B.(1939) Reliability of mental tests. **Br. J. Psychol.** 29, 267-287.
- [29] Johnson, N.L. & Leone, F.C. (1964) **Statistics and Experimental Design: In Engineering and The Physical Sciences, Volume II**. New York: John Wiley & Sons.
- [30] Kempthorne, O.(1957) **An Introduction to Genetic Statistics** New York: John Wiley & Sons.
- [31] Khuri, A.I. & Sahai, H. (1985) Variance components analysis: a selective literature survey. **International Statistical Review** 53, 3, 279-300.
- [32] Morettin, P.A. & Bussab, W. O.(1984) **Estatística Básica**. São Paulo: Atual.

- [33] Neter, J., Wasserman, W. & Kutner, M. H.(1990) **Applied Linear Statistical Models: regression, analysis of variance and experimental designs**. Homewood: Richard D. Irwin.
- [34] Neyman, J. Iwaszkiewicz, K. & Kolodziejczyk, S.T.(1935) Statistical problem in agricultural experimentation. **J. R. Statist. Soc. Suppl. 2**, 107-154.
- [35] Peres, C.A.(1981) Testing the effect of blocking in a randomized complete block design (RCBD). **Commun. Statist. - Theor. Meth. A(10) 23**, 2447-2459.
- [36] Peres, C.A.(1990) Optimum allocation for random effects model. **XVth International Biometric Conference**, Budapest.
- [37] Peres, C.A. & Saldiva, C.D.(1982) **Planejamento de Experimentos**. São Paulo: IME-USP.
- [38] Rao, C.R.(1973) **Linear Statistical Inference and Its Applications**. New York: John Wiley & Sons.
- [39] Rao, C.R. and Kleffe, J.(1988) **Estimation of Variance Components and Applications**. Amsterdam: North-Holland.
- [40] Sahai, H.(1979) A bibliography on variance components. **International Statistical Review 47**, 177-222.
- [41] Sahai, H., Khuri, A.I. & Kapadia, C.H.(1985) A second bibliography on variance components. **Communication Statistics A 14**, 63-115.
- [42] Scheffé, H.(1956) Alternative models for the analysis of variance. **Ann. Math. Statist. 27**, 251- 271.
- [43] Scheffé, H.(1959) **The Analysis of Variance**. New York: John Wiley & Sons.
- [44] Searle, S.R.(1971) **Linear Models**. New York: John Wiley & Sons.

- [45] Smith, C.A.B.(1956) On the estimation of intraclass correlation. **Ann. Hum. Gen.** **21**, 363-373.
- [46] Snedecor, G.W. & Cochran, W.G.(1967) **Statistical Methods**. Ames: The Iowa State University.
- [47] Swallow, W.H. & Monahan, J.F.(1984) Monte-Carlo comparison of ANOVA, MIVQUE, REML, and ML of estimators of variance components. **Technometrics** **26**,47-57.
- [48] Tippet, L.H.C. (1931) **The Methods of Statistics**. London: Williams and Norgate.
- [49] Winsor, C.P. & Clarke, G.L. (1940) Statistical study of variation in the catch of plankton nets. **J. Marine Res.** **3**,1-34.
- [50] Yates, F. & Zecopanay, I. (1935) The estimation of the efficiency of sampling with special reference to sampling for yield in cereal experiments. **J. Agric. Sci.** **25**, 545-577.