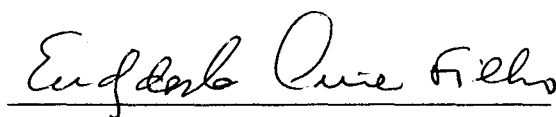


**"O MÉTODO DE QUADRADOS MÍNIMOS PONDERADOS NA
ANÁLISE DE TABELAS DE CONTINGÊNCIA: APLICAÇÕES
EM PESQUISA AGRÍCOLA."**

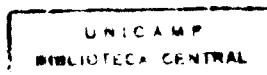
Este exemplar corresponde à redacção final da
tese devidamente corrigida e defendida pela Sra.
ALINE DE HOLANDA NUNES MAIA e
aprovada pela comissão julgadora.

Campinas, 02 de setembro de 1994.



Prof. Dr. Euclides Custódio de Lima Filho

Dissertação apresentada ao Instituto de
Matemática Estatística e Ciência da
Computação, UNICAMP, como requisito parcial
para obtenção do Título de MESTRE em
Estatística.



AGRADECIMENTOS

Agradeço:

- À Empresa Brasileira de Pesquisa Agropecuária pelo apoio financeiro;
- Aos professores do IMECC, pelo estímulo, em especial ao Prof. Euclides;
- Aos colegas, pelo companheirismo e solidariedade nos momentos difíceis.

A Luiz Antônio, pelo carinho e apoio em todos os momentos.

Índice

1	Introdução.	4
2	O modelo linear	7
2.1	Modelo de coordenadas livres versus modelo de coordenadas fixas.	8
2.1.1	Modelo para uma amostra de uma única população.	11
2.1.2	Análise de variância com um fator.	12
2.1.3	Análise de regressão múltipla	14
2.2	Avaliação da qualidade do modelo.	14
2.2.1	Coefficiente de determinação.	15
2.2.2	Análise de resíduos.	15
2.3	Estimação de parâmetros.	17
2.4	Testes de hipóteses lineares sobre μ	19
3	Análise de tabelas de contingência via método de quadrados mínimos ponderados.	22
3.1	Funções de resposta.	23
3.1.1	Escolha da função de resposta.	25
3.1.2	Distribuição assintótica de F	27
3.2	O modelo linear geral	29
3.3	Ajuste do modelo	30
3.4	Análise de resíduos	31
3.5	Método de quadrados mínimos ponderados.	33
3.5.1	Estimação do vetor de parâmetros do modelo.	34
3.5.2	Estimação da função de resposta.	35

3.6	Propriedades dos estimadores de quadrados mínimos ponderados.	35
3.7	Testes de hipóteses lineares	37
4	Modelagem probabilística das tabelas de contingência.	40
4.1	Dados modelados pela distribuição produto de multinomiais,	40
4.2	Dados modelados pela distribuição produto de Poisson.	45
5	Exemplos.	51
5.1	Exemplo 1. Associação entre peso do fruto e escurecimento interno em abacaxi Smooth Cayenne.	51
5.1.1	Análise dos dados considerando a estrutura I.	53
5.1.2	Análise dos dados considerando a estrutura II.	65
5.2	Exemplo 2. Comparação de testes de vigor para prever emergência em sorgo.	88
5.2.1	Esperança e covariância de p	91
5.2.2	Escolha da função de resposta.	91
5.2.3	Distribuição assintótica de $F(p)$	95
5.2.4	Análise dos resultados.	97
5.3	Exemplo 3. Efeito do substrato no enraizamento de alporques de urucuzeiro.	101
5.3.1	Esperança e covariância de Y	103
5.3.2	Escolha da função de resposta	104
5.3.3	Definição do modelo e formulação de hipóteses.	105
5.3.4	Análise dos resultados.	106
6	Considerações finais.	109
A	Algumas definições e teoremas em álgebra linear.	110
A.1	Espaços vetoriais, subespaços e bases.	110
A.2	Transformações lineares.	113
A.3	Projetores.	116

B	Propriedades assintóticas de estimadores e estatísticas de teste.	120
B.1	Algumas propriedades assintóticas de estimadores.	120
B.1.1	O método δ para o cálculo de distribuições assintóticas.	123
B.2	Estatísticas de teste com distribuição assintótica χ^2	126
C	Métodos computacionais.	128

Capítulo 1

Introdução.

Em muitas pesquisas agrícolas, as variáveis resposta de interesse são categorizadas, isto é, mensuradas em escala nominal ou ordinal. Por exemplo, o número de espigas de milho por planta, número de grãos por vagem, tamanho de bulbos (pequeno, médio e grande), estado de conservação de frutos (bom, regular, péssimo), entre outras, pertencem a essa classe de variáveis.

Nos experimentos agrícolas, quando as variáveis resposta são categóricas, mas assumem valores numéricos, a técnica de análise usualmente empregada é buscar transformações dessas variáveis de modo satisfazer os pressupostos da análise de variância do modelo linear comum e realizar comparações múltiplas entre as médias dos tratamentos.

O objetivo desse trabalho é apresentar e discutir o método de quadrados mínimos ponderados (QMP), também conhecido como método GSK, como alternativa para análise de dados dessa natureza. Nessa discussão, é enfatizada a representação geométrica das estatísticas de interesse buscando facilitar sua interpretação e as provas de teoremas relacionados.

As técnicas estatísticas para análise de dados categóricos foram desenvolvidas a partir dos anos 60. Nessa década, muitos trabalhos foram publicados discutindo a abordagem que tornou-se conhecida como modelos log-lineares. Bishop *et al.* (1975), Fienberg (1980), Haberman (1978, 1979) publicaram livros onde estão resumidos muitos dos trabalhos baseados na teoria de máxima verossimilhança, desenvolvidos na década anterior.

Grizzle, Starmer e Koch (1969) propuseram uma abordagem alternativa para a análise

de dados categorizados com base na teoria de modelos lineares e quadrados mínimos ponderados, conhecida como método GSK. Nesse trabalho, é descrito um método não iterativo para ajuste de um modelo linear a um conjunto de dados com distribuição produto de multinomiais, baseado nas estatísticas de Wald (1943) e Neyman (1949). O interesse é descrever a variabilidade entre funções do parâmetro π que indexa a distribuição dos dados. Essas funções devem ser tais que as derivadas até ordem dois com respeito às componentes de π existam e a matriz de derivadas de primeira ordem tenha posto linha completo. Em particular, são analisados casos onde as funções de interesse são lineares ou logarítmicas .

Forthofer & Koch (1973) generalizam a aplicação do método GSK para o caso de combinações de funções lineares, exponenciais e logarítmicas. Muitos dos trabalhos sobre a abordagem GSK estão resumidos em Forthofer & Lehnen (1981). Eles destacam a flexibilidade do método de QMP e apresentam uma grande variedade de aplicações.

Freeman (1982), ressalta que o método de QMP é computacionalmente mais simples que o de máxima verossimilhança e permite a estimação de funções complexas de variáveis categóricas.

Koch *et al* (1983), apresentam as abordagens de máxima verossimilhança, quadrados mínimos ponderados e testes aleatorizados na análise de dados categóricos. Esses autores destacam que quando os dados analisados têm distribuição de Poisson ou multinomial, o método de Q.M.P apresenta algumas características importantes:

1. Os estimadores de Q.M.P. pertencem à classe dos estimadores BAN (best asymptotically normal), isto é, têm mínima variância assintótica entre todos os estimadores não viciados do parâmetro de interesse e são assintoticamente normais;
2. O estimador de Q.M.P. do vetor de parâmetros é assintoticamente equivalente ao estimador de máxima verossimilhança, no sentido de que possui a mesma distribuição aproximada quando os tamanhos das amostras utilizadas para o cálculo das componentes da função de interesse são suficientemente grandes;

3. A estatística de Wald para teste do ajuste do modelo proposto e as estatísticas dos testes de hipóteses lineares são assintoticamente equivalentes às estatísticas correspondentes baseadas na razão de verossimilhança;

Em todas as publicações anteriormente referidas, os exemplos objeto de análise provêm quase exclusivamente da área de ciências médicas ou sociais. Verifica-se que pouca ênfase é dada à aplicação desses métodos na pesquisa agrícola.

Essa dissertação constitui-se de cinco capítulos e dois apêndices, cujo conteúdo é resumido a seguir. No capítulo 2, é feita uma revisão sobre o modelo linear comum incluindo avaliação da qualidade do modelo, estimação de parâmetros pelo método de quadrados mínimos ordinários e testes de hipóteses lineares. Além da postulação usual de coordenadas fixas, é apresentada a postulação de coordenadas livres conforme descrita em Arnold (1981). No capítulo 3 são caracterizadas situações em que o método QMP é utilizado, a modelagem da esperança assintótica de funções de estimadores de parâmetros de interesse, descreve-se o método de estimação, as propriedades assintóticas dos estimadores resultantes e testes de hipóteses lineares sobre o vetor de parâmetros do modelo proposto. No capítulo 4 caracterizam-se situações em que os dados são adequadamente modelados pelas distribuições de Poisson e multinomial e obtidos estimadores consistentes das matrizes de covariância assintóticas de funções dos parâmetros que indexam tais distribuições. Análises detalhadas de aplicações na pesquisa agrícola são apresentadas no capítulo 5. No apêndice A, são apresentados definições e teoremas em álgebra linear e no apêndice B, definições e teoremas relacionados com propriedades assintóticas de estimadores e estatísticas de teste. No apêndice C constam os programas utilizados para análise dos exemplos do capítulo 5. Os teoremas utilizados nessa dissertação serão identificados no texto através da letra do apêndice seguida do número do teorema.

Capítulo 2

O modelo linear

Um modelo estatístico resulta de técnicas que buscam expressar de forma quantitativa a relação entre variáveis resposta e variáveis preditoras. Quando o interesse concentra-se na modelagem de uma única resposta, o modelo é dito univariado. Nesse caso a informação sobre a natureza dessa relação é obtida através de um vetor de observações $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ que corresponde a uma amostra aleatória de populações conceituais de observações da variável resposta.

Um modelo estatístico é dito linear (nos parâmetros) se postula que a esperança de cada observação Y_i pode ser escrita como uma combinação linear de p parâmetros desconhecidos $\beta_1, \beta_2, \dots, \beta_p$. Arnold(1981), Joshi(1987), Searle(1971), entre outros, caracterizam como *modelo linear* aquele que além de ser uma função linear dos parâmetros, satisfaz às seguintes pressuposições:

- i) As variáveis Y_i têm todas a mesma variância σ^2 (desconhecida) e são não correlacionadas, isto é, a matriz de covariâncias de \mathbf{Y} é da forma $\sigma^2\mathbf{I}$, ($\sigma^2 > 0$);
- ii) As variáveis Y_i têm distribuição normal .

Quando o modelo é linear nos parâmetros mas alguma das pressuposições acima não se cumpre o modelo é dito *modelo linear geral*.

É importante ressaltar que na maioria dos casos, as verdadeiras relações entre as variáveis envolvidas nos fenômenos que se buscam modelar não são lineares. Usam-se modelos lineares porque freqüentemente eles representam uma boa aproximação dessas relações. Além disso, a natureza aditiva dos modelos lineares torna-os mais fáceis de

trabalhar matematicamente que os modelos não lineares.

Um modelo linear pode ser postulado de diferentes formas. Na seção 2.1. serão discutidas as postulações de *coordenadas livres* e de *coordenadas fixas*, descritas em Arnold (1981).

A concordância entre o modelo proposto e os dados observados é avaliada através de estatísticas de *ajuste do modelo* e da análise de *resíduos*. No modelo linear comum, usa-se como estatística do ajuste o *coeficiente de determinação*. Esse coeficiente mede a porcentagem da variabilidade total de um particular conjunto de dados que é explicada pelo modelo proposto. Os resíduos são a diferença entre os valores *observados* e *preditos* pelo modelo. A *análise dos resíduos* busca detectar o não cumprimento das pressuposições (i) e (ii).

Quando a variabilidade da resposta que se deseja modelar é adequadamente descrita através do modelo proposto e não há evidência suficiente de violação das pressuposições, as etapas subseqüentes da análise são a *estimação de parâmetros* desconhecidos e a realização de *testes de hipóteses* sobre eles. A adequação do modelo proposto (ajuste e análise de resíduos), estimação de parâmetros e testes de hipóteses no modelo linear serão objeto das seções 2.2 e 2.3 e 2.4, respectivamente.

Nesse capítulo, as estatísticas de interesse no modelo linear são expressas como projeções ou normas de projeções com base na postulação do modelo de coordenadas livres. Essa abordagem busca facilitar a interpretação de estatísticas de interesse e provas de teoremas.

2.1 Modelo de coordenadas livres versus modelo de coordenadas fixas.

Seja V um subespaço do IR_n com dimensão p e $Y = (Y_1, Y_2, \dots, Y_n)$ um vetor aleatório observável. Considere que pressuposições (i) e (ii) valem.

Denote $E(Y) = \mu$. Um modelo linear pode ser então postulado afirmando-se que

$\mu \in V$. Como $\dim(V) = p$, qualquer matriz \mathbf{B} cujas colunas constituem uma base de V , tem posto-coluna completo igual a p . Sejam $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p$ as colunas de \mathbf{B} . Então, se $\mu \in V$, μ pode ser escrito como combinação linear das colunas $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_p$ de qualquer base \mathbf{B} , isto é, existem escalares $\gamma_1, \gamma_2, \dots, \gamma_p$ não todos nulos tais que:

$$\mu = \gamma_1 \mathbf{b}_1 + \gamma_2 \mathbf{b}_2 + \dots + \gamma_p \mathbf{b}_p$$

Dessa forma, postular que

$$\mathbf{Y} \sim \mathbf{N}(\mu, \sigma^2 \mathbf{I}), \mu \in V; \sigma^2 > 0 \quad (2.1)$$

corresponde ao modelo

$$\mu = \mathbf{B}\gamma$$

pois:

$$\begin{aligned} \mu &= \gamma_1 \begin{pmatrix} b_{11} \\ \cdot \\ \cdot \\ \cdot \\ b_{n1} \end{pmatrix} + \gamma_2 \begin{pmatrix} b_{12} \\ \cdot \\ \cdot \\ \cdot \\ b_{n2} \end{pmatrix} + \dots + \gamma_p \begin{pmatrix} b_{1p} \\ \cdot \\ \cdot \\ \cdot \\ b_{np} \end{pmatrix} = \\ &= \begin{pmatrix} b_{11}\gamma_1 + \dots + b_{1p}\gamma_p \\ \cdot \\ \cdot \\ \cdot \\ b_{n1}\gamma_1 + \dots + b_{np}\gamma_p \end{pmatrix} = \end{aligned}$$

$$= \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \cdot \\ \cdot \\ \cdot \\ \gamma_p \end{pmatrix} = \mathbf{B}\boldsymbol{\gamma}$$

onde $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_p)$ é a parametrização correspondente à base \mathbf{B} . Os parâmetros $\gamma_1, \gamma_2, \dots, \gamma_p$ para uma particular escolha de base \mathbf{B} são as coordenadas do vetor $\boldsymbol{\mu}$ com relação à base \mathbf{B} .

Como não foi escolhida uma base, as coordenadas do vetor de esperanças ($E(\mathbf{Y}) = \boldsymbol{\mu}$) com relação à base não foram fixadas e o modelo assim postulado é dito *modelo de coordenadas livres*.

Seja \mathbf{X} uma base de V . Então \mathbf{X} é $(n \times p)$ de posto p , $\mathbf{X}'\mathbf{X}$ é inversível e existe um único $\boldsymbol{\beta} \in R_p$ tal que $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ (Teorema A2). Dessa forma,

$$\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta} \Leftrightarrow \mathbf{X}'\boldsymbol{\mu} = \mathbf{X}'\mathbf{X}\boldsymbol{\beta} \Leftrightarrow \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\mu}. \quad (2.2)$$

Escolhida uma base para V , representada pelas p colunas $\mathbf{x}_j = (x_{1j}, x_{2j}, \dots, x_{nj})$ da matriz \mathbf{X} , o modelo passa a ser postulado da seguinte forma:

$$\begin{pmatrix} E(Y_1) \\ \cdot \\ \cdot \\ \cdot \\ E(Y_n) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \cdot \\ \cdot \\ \cdot \\ \mu_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_2 \end{pmatrix}$$

onde valem as pressuposições (i) e (ii). Uma forma resumida dessa mesma postulação é feita usando notação matricial:

$$\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}), \quad \sigma^2 > 0. \quad (2.3)$$

O modelo definido acima é dito *modelo de coordenadas fixas*.

Arnold (1981), enumera as seguintes razões para o uso do modelo de coordenadas livres:

(1) Não é necessário escolher uma *base* para o subespaço V . Esse subespaço pode ser descrito por uma de suas bases ou por um conjunto de *restrições lineares* que os elementos do subespaço devem satisfazer. A segunda abordagem parece mais razoável para problemas de análise de variância, onde freqüentemente não existe uma base natural para o fenômeno;

(2) Na postulação do modelo livre de coordenadas, a maior parte das estatísticas de interesse podem ser interpretadas como *projeções* ou *normas de projeções*. Como essas projeções não dependem da base, geralmente, toma-se uma *base ortogonal*, o que simplifica as provas;

(3) As fórmulas das estatísticas são muito mais fáceis de interpretar quando expressas como projeções que quando expressas como matrizes;

(4) Em muitos problemas particulares de análise de variância, as fórmulas são mais fáceis de derivar usando propriedades de projeções que expressões matriciais equivalentes.

A seguir serão apresentados alguns exemplos das duas formulações do modelo para cada problema específico.

2.1.1 Modelo para uma amostra de uma única população.

Sejam Y_1, Y_2, \dots, Y_n variáveis aleatórias independentes com distribuição comum $N_1(\theta, \sigma^2)$, $\theta \in R$, $\sigma^2 \in (0, \infty)$. Usando notação matricial:

$$\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I})$$

onde $\boldsymbol{\mu} = (\theta, \theta, \dots, \theta)$ e $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$

Essa postulação corresponde à idéia de que as componentes do vetor observado \mathbf{Y} são oriundas de uma mesma população e conseqüentemente o vetor de esperanças $\boldsymbol{\mu}$ pertence ao subespaço dos vetores de R^n onde todas as componentes são iguais a θ para algum

$\theta \in R$. Seja V_1 esse subespaço. Então, uma versão equivalente desse modelo é:

$$\mathbf{Y} \sim N_n(\mu, \sigma^2 \mathbf{I}); \mu \in V_1; \sigma^2 \in (0, \infty) \quad (2.4)$$

O modelo descrito acima corresponde à versão livre de coordenadas .

Como o subespaço V_1 tem dimensão um, qualquer vetor pertencente a esse subespaço constitui uma base.

Seja $\mathbf{X}_{n \times 1} = (1, 1, \dots, 1)$. O modelo de coordenadas fixas equivalente a (2.4) para essa escolha de base é:

$$\mu = \mathbf{X}\beta$$

e nesse caso , $\beta_{1 \times 1} = \theta$ é um escalar e representa a média populacional.

2.1.2 Análise de variância com um fator.

Deseja-se avaliar o efeito de um fator (tratamento) com k níveis T_1, T_2, \dots, T_k sobre alguma resposta de interesse. Em cada subpopulação é retirada uma amostra Y_1, Y_2, \dots, Y_n , representativa da população conceitual correspondente ao tratamento i ($i = 1, 2, \dots, k$). Considera-se que $Y_{ij} \sim N(\theta_i, \sigma^2)$, $\theta_i \in R$, $\sigma^2 \in (0, \infty)$, ($j = 1, 2, \dots, n_i$). Na forma matricial o modelo é expresso como:

$$\mathbf{Y} \sim N_n(\mu, \sigma^2 \mathbf{I})$$

onde $\mu = (\theta_1, \dots, \theta_1, \dots, \theta_k, \dots, \theta_k)$ e $\mathbf{Y} = (Y_{11}, \dots, Y_{1n_1}, \dots, Y_{k1}, \dots, Y_{kn_k})$ e $n = \sum n_k$.

Seja V_2 um subespaço k -dimensional do R^n onde os n_1 primeiros elementos são iguais a θ_1 , os n_2 elementos seguintes iguais a θ_2 , ..., os n_k últimos elementos são iguais a θ_k , para alguma k -upla $(\theta_1, \theta_2, \dots, \theta_k) \in R^k$. Então a formulação livre de coordenadas para o modelo é:

$$\mathbf{Y} \sim N_n(\mu, \sigma^2 \mathbf{I}); \mu \in V_2; \sigma^2 \in (0, \infty) \quad (2.5)$$

Seja a matriz X abaixo:

$$X_{n \times k} = \begin{bmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 1 & -1 & -1 & -1 & \dots & -1 & -1 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 1 & -1 & -1 & -1 & \dots & -1 & -1 \end{bmatrix}$$

X é $n \times k$ de posto k e portanto possui k colunas linearmente independentes que constituem uma base de V_2 .

O modelo de coordenadas fixas correspondente a (2.5), para essa particular escolha de base é:

$$\mu = X\beta$$

onde as componentes do vetor de parâmetros $\beta = (\beta_0, \beta_1, \dots, \beta_{k-1})$ têm a seguinte interpretação:

β_0 = média geral;

β_i = efeito diferencial do tratamento i , com relação à média ($i = 1, 2, \dots, k - 1$).

A cada base (matriz) escolhida corresponde uma parametrização e vice-versa. A parametrização deve ser tal que X tenha posto completo e que os parâmetros sejam interpretáveis.

2.1.3 Análise de regressão múltipla

Deseja-se explicar a variabilidade de uma resposta Y através de um conjunto de variáveis preditoras X_1, X_2, \dots, X_k (conhecidas). Sejam Y_1, Y_2, \dots, Y_n uma amostra aleatória com $Y_i \sim N_1(\sum_{j=1}^k x_{ij}\beta_j, \sigma^2)$ com $\beta_j \in R, \sigma^2 > 0$. Nesse caso existe uma base natural para o modelo que é então postulado, da seguinte forma

$$\mu = \mathbf{X}\beta$$

onde $\mathbf{X} = [x_{ij}]$ é uma matriz $n \times k$ de constantes conhecidas correspondentes aos níveis das variáveis preditoras X_1, X_2, \dots, X_k e β é vetor $k \times 1$ de parâmetros desconhecidos do modelo. Apesar da existência dessa base natural podemos escolher outra base \mathbf{B} qualquer tal que o espaço-coluna de \mathbf{B} , denotado por $C(\mathbf{B})$ é igual ao espaço-coluna de $C(\mathbf{X})$, sem afetar o ajuste do modelo, os valores preditos e os resíduos. Tal fato é justificado pelo Teorema A15 e será explicado nas seções subsequentes. Muitas vezes, para resolver problemas de *mau condicionamento* da base natural \mathbf{X} , usa-se uma base ortonormal do $C(\mathbf{X})$. Para maiores detalhes sobre condicionamento da matriz do modelo e outros *diagnósticos de colinearidade*, ver Montgomery & Peck, (1982).

2.2 Avaliação da qualidade do modelo.

A postulação de um modelo estatístico inclui, além do modelo matemático proposto, pressuposições sobre a distribuição dos dados. Portanto, a qualidade de um modelo estatístico relaciona-se com o ajuste dos dados ao modelo matemático proposto e o cumprimento das pressuposições. Como medida geral de ajuste, usa-se o *coeficiente de determinação*. Evidências sobre o possível não cumprimento das pressuposições são investigadas através da *análise de resíduos*.

2.2.1 Coeficiente de determinação.

O coeficiente de determinação (R^2) é uma medida resumo que representa a percentagem de variabilidade de um conjunto de dados que é explicada pelo modelo proposto. Geometricamente, corresponde ao cosseno do ângulo (α) entre o vetor de observações \mathbf{Y} e sua projeção ortogonal em V , $\mathbf{P}_V^* \mathbf{Y}$:

$$R^2 = \cos(\alpha) = \frac{\|\mathbf{P}_V^* \mathbf{Y}\|^2}{\|\mathbf{Y}\|^2} \quad (2.6)$$

$\|\mathbf{P}_V^* \mathbf{Y}\|^2$ é a variabilidade explicada pelo modelo (Soma de quadrados devida ao modelo) e $\|\mathbf{Y}\|^2$, a variabilidade total (Soma de quadrados total).

Como $0 \leq \|\mathbf{P}_V^* \mathbf{Y}\| \leq \|\mathbf{Y}\|$ (Teorema A15-d), o coeficiente de determinação definido em (2.6) pode assumir valores entre 0 e 1. Observe que o coeficiente de determinação independe da base escolhida.

Como medida de ajuste global, o coeficiente de determinação não dá evidência sobre o não cumprimento das pressuposições do modelo linear (normalidade, homocedasticidade e independência).

2.2.2 Análise de resíduos.

Seja o modelo ajustado

$$\mathbf{Y} = \mathbf{X}\hat{\beta} + \mathbf{e}$$

correspondente ao modelo de coordenadas fixas definido em (2.3). $\hat{\beta}$ é o *estimador de quadrados mínimos* de β e \mathbf{e} o vetor de resíduos.

Quando o vetor de parâmetros do modelo é estimado pelo método de quadrados mínimos, o vetor de resíduos (\mathbf{e}) é ortogonal ao vetor de estimativas ($\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$). Assim, o vetor de observações \mathbf{Y} pode ser decomposto na soma direta ortogonal de $\hat{\mathbf{Y}} = \mathbf{P}_V^* \mathbf{Y}$ e \mathbf{e} , isto é,

$$\mathbf{Y} = \hat{\mathbf{Y}} \oplus \mathbf{e}.$$

Logo, o vetor de resíduos é o complemento ortogonal de \mathbf{Y} em IR^n e pode ser expresso como projeção ortogonal de \mathbf{Y} em V^\perp . Pelo teorema (A15-a), a matriz do projetor ortogonal de \mathbf{Y} em V^\perp é $(\mathbf{I} - \mathbf{P}_V^*)$. Então

$$\mathbf{e} = (\mathbf{I} - \mathbf{P}_V^*)\mathbf{Y}$$

O vetor de resíduos ($\mathbf{e} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$) é um vetor aleatório e sua observação fornece informação sobre os erros. A análise de resíduos busca evidências sobre o não cumprimento das pressuposições definidas em (2.3).

As pressuposições sobre o vetor de observações que caracterizam o modelo linear (comum) podem ser resumidas em

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{I}) \quad (2.7)$$

Considerando (2.7) válido,

$$\mathbf{E}(\boldsymbol{\epsilon}) = \mathbf{E}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{E}(\mathbf{Y}) - \mathbf{E}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}\boldsymbol{\beta} - \mathbf{X}\boldsymbol{\beta} = \mathbf{0}$$

e

$$\mathbf{Var}(\boldsymbol{\epsilon}) = \mathbf{Var}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{Var}(\mathbf{Y}) = \sigma^2 \mathbf{I}.$$

Então (2.7) pode ser reescrito como:

$$\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}). \quad (2.8)$$

Montgomery & Peck (1982) apresentam diversos procedimentos gráficos baseados na análise de resíduos, úteis para detectar possíveis violações das pressuposições, entre eles:

1. Gráfico das probabilidades normais (*Normal probability plot*): graficam-se os resíduos num papel probabilístico. Se o gráfico resultante é próximo de uma reta, é uma indicação de que a pressuposição de normalidade dos erros não foi violada.
2. Gráfico dos resíduos versus valores preditos: pode indicar o não cumprimento da pressuposição de homocedasticidade (igualdade das variâncias dos erros).
3. Gráfico dos resíduos versus variáveis preditoras: pode indicar ausência de homocedasticidade, relação não linear entre preditoras e resposta e/ou inclusão de possíveis novas preditoras.
4. Resíduos versus variáveis potenciais não incluídas no modelo: indica se a variabilidade não explicada pelo modelo atual (resíduos) tem relação com possíveis novas variáveis preditoras.

2.3 Estimação de parâmetros.

No modelo de coordenadas livres descrito na seção 2.1, postula-se que $\mathbf{E}(\mathbf{Y}) = \boldsymbol{\mu}$ pertence a um subespaço $V \subset \mathbb{R}^n$. Seja $\mathbf{B}_{n \times p}$ uma base qualquer desse subespaço, de modo que $\boldsymbol{\mu}$ pertence a V se e somente se $\boldsymbol{\mu} = \mathbf{B}\boldsymbol{\gamma}$ para algum $\boldsymbol{\gamma} \in \mathbb{R}^p$. Note que $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_p)$ é a parametrização correspondente à escolha da base \mathbf{B} de V .

O vetor $\boldsymbol{\epsilon} = \mathbf{Y} - \mathbf{B}\boldsymbol{\gamma}$ é chamado vetor de erros e é suposto ter distribuição normal com $\mathbf{E}(\boldsymbol{\epsilon}) = \mathbf{o}$ e $\text{Var}(\boldsymbol{\epsilon}) = \sigma^2 \mathbf{I}$. Resulta da diferença entre o vetor de observações \mathbf{Y} e seu valor esperado sob o modelo, $\mathbf{E}(\mathbf{Y}) = \mathbf{B}\boldsymbol{\gamma} \in C(\mathbf{B})$. O quadrado da norma euclidiana do vetor de erros

$$(\mathbf{Y} - \mathbf{B}\boldsymbol{\gamma})'(\mathbf{Y} - \mathbf{B}\boldsymbol{\gamma}) \quad (2.9)$$

é chamada soma de quadrados dos erros (SQE) e representa a distância euclidiana entre o vetor aleatório \mathbf{Y} e $\mathbf{B}\boldsymbol{\gamma}$.

O método de *quadrados mínimos* é um método de estimação de parâmetros que consiste em encontrar $\boldsymbol{\gamma} \in \mathbb{R}^p$ tal que a expressão (2.9) seja minimizada.

Seja $\mathbf{y} = (y_1, y_2, \dots, y_n)$ uma observação de \mathbf{Y} e \mathbf{P}_V^* o projetor ortogonal do \mathbb{R}^n no subespaço V . O vetor de erros correspondente à observação \mathbf{y} é:

$$(\mathbf{y} - \boldsymbol{\mu})'(\mathbf{y} - \boldsymbol{\mu})$$

No teorema (A16), é mostrado que a distância entre \mathbf{y} e sua projeção ortogonal em V , $\mathbf{P}_V^* \mathbf{y}$ satisfaz:

$$(\mathbf{y} - \mathbf{P}_V^* \mathbf{y})'(\mathbf{y} - \mathbf{P}_V^* \mathbf{y}) = \min_{\boldsymbol{\mu} \in V} \{(\mathbf{y} - \boldsymbol{\mu})'(\mathbf{y} - \boldsymbol{\mu})\}.$$

Dessa forma, a expressão (2.9) é minimizada quando $\boldsymbol{\gamma}$ é tal que $\mathbf{P}_V^* \mathbf{y} = \mathbf{B}\boldsymbol{\gamma}$. Pelo Teorema A14, a matriz do projetor \mathbf{P}_V^* (também denotada por \mathbf{P}_V^*) é expressa como:

$$\mathbf{P}_V^* = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'.$$

Seja \mathbf{X} uma base de V e $\boldsymbol{\beta}$ a parametrização correspondente a essa escolha de base. Conforme o Teorema A15, a matriz do projetor ortogonal independe da base, de modo que, fixada a base \mathbf{X} de V ,

$$\mathbf{P}_V^* = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

Fazendo $\mathbf{B} = \mathbf{X}$ e $\boldsymbol{\gamma} = \boldsymbol{\beta}$ a expressão (2.9), resulta em

$$(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \tag{2.10}$$

e $\mathbf{X}\boldsymbol{\beta}$ tal que (2.10) é mínima e deve satisfazer

$$\mathbf{P}_V^* \mathbf{y} = \mathbf{X}\boldsymbol{\beta} \Rightarrow \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{X}\boldsymbol{\beta}. \tag{2.11}$$

Pre-multiplicando as expressões da direita e da esquerda em (2.11) por \mathbf{X}' , obtemos

$$\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{2.12}$$

De (2.12) resulta o sistema de equações

$$\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{y}. \quad (2.13)$$

Esse sistema é conhecido como *equações normais*. Como \mathbf{X} é uma base de V , \mathbf{X} tem posto coluna completo (posto de \mathbf{X} igual a p), $\mathbf{X}'\mathbf{X}$ é uma matriz quadrada de ordem p não singular e as equações normais têm uma única solução

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \quad (2.14)$$

$\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)$ é chamado *estimador de quadrados mínimos* de β .

Quando a matriz do modelo não tem posto-coluna completo, o sistema (2.14) tem infinitas soluções e não se pode falar de estimador de mínimos quadrados de β . Qualquer solução $\hat{\beta}$ das equações normais satisfaz

$$(\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta}) = \min_{\beta \in \mathbb{R}^p} [(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta)] \quad (2.15)$$

e o estimador de quadrados mínimos de $\mathbf{X}\beta$ é $\mathbf{X}\hat{\beta}$.

2.4 Testes de hipóteses lineares sobre μ .

Considere a postulação de coordenadas livres

$$\mathbf{Y} \sim \mathbf{N}(\mu, \sigma^2\mathbf{I}), \text{ onde } \mu \in V, \sigma^2 > 0$$

definida em (2.1).

Uma hipótese H_0 sobre o vetor de esperanças μ é dita uma *hipótese linear* se pode ser representada por um conjunto de *contrastés lineares*

$$H_0 : \begin{cases} c_{11}\mu_1 + \dots + c_{1n}\mu_n = 0 \\ c_{21}\mu_1 + \dots + c_{2n}\mu_n = 0 \\ \vdots \\ c_{k1}\mu_1 + \dots + c_{kn}\mu_n = 0 \end{cases}$$

onde $\mathbf{C} = [c_{ij}]$ é uma matriz $k \times n$ de posto k . Usualmente, H_0 é apresentada de forma resumida como

$$H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{0}.$$

De modo alternativo, H_0 pode ser postulada afirmando que $\boldsymbol{\mu}$ pertence a um subespaço W de V , isto é

$$\boldsymbol{\mu} \in W \subset V \subset R^n$$

onde W é um subespaço de V tal que W é ortogonal ao espaço-linha de \mathbf{C} , denotado por $L(\mathbf{C})$.

Como as dimensões de V e $L(\mathbf{C})$ são p e k , respectivamente, e

$$V = W \oplus L(\mathbf{C})$$

a dimensão de W é $p - k$ e a dimensão de $W^\perp \cap V$ é k .

Sob H_0 , $\boldsymbol{\mu} \in W$. Quando o vetor de observações \mathbf{Y} está distante do subespaço W é uma indicação de que a hipótese é falsa. Uma maneira de quantificar essa proximidade é calcular a distância euclidiana entre \mathbf{Y} e sua projeção ortogonal em W ,

$$\|\mathbf{Y} - \mathbf{P}_W^* \mathbf{Y}\|^2 = \|(\mathbf{I} - \mathbf{P}_W^*) \mathbf{Y}\|^2 = \|\mathbf{P}_{W^\perp \cap V}^* \mathbf{Y}\|^2$$

e compará-la com o quadrado da norma euclidiana do vetor de resíduos (soma de quadrados de resíduos) do modelo (2.6),

$$\|\mathbf{Y} - \mathbf{P}_V^* \mathbf{Y}\|^2 = \|(\mathbf{I} - \mathbf{P}_V^*) \mathbf{Y}\|^2 = \|\mathbf{P}_{V^\perp}^* \mathbf{Y}\|^2$$

ponderando pelos respectivos graus de liberdade.

O teste geralmente utilizado para hipóteses dessa natureza em modelo lineares é o teste F generalizado, cuja estatística é

$$F = \frac{\|P_{W^\perp \cap V}^* Y\|^2 / k}{\|P_{V^\perp}^* Y\|^2 / (n - p)}.$$

Sob H_0 , F tem distribuição F de Snedecor com k graus de liberdade do numerador e $n - p$ graus de liberdade do denominador.

Para decidir sobre a rejeição ou não de H_0 , calcula-se o valor de F e o correspondente valor p .

Capítulo 3

Análise de tabelas de contingência via método de quadrados mínimos ponderados.

O método de quadrados mínimos ponderados é usado para estimar parâmetros em *modelos lineares* que descrevem a variação entre estatísticas (*funções*) de interesse na análise de *dados categóricos*. Essa abordagem foi proposta por Grizzle, Starmer e Koch (1969) e é conhecida como método GSK. O método de estimação utilizado baseia-se em estatísticas propostas por Wald (1943) e Neyman (1949).

O conjunto de observações de dados categorizados objeto de análise pode ser resumido numa tabela de contingência, que é uma representação resultante da classificação cruzada de duas ou mais variáveis categorizadas. Geralmente, as linhas da tabela correspondem a subpopulações (definidas pelos níveis das variáveis preditoras) e as colunas correspondem a níveis da variável resposta. Em alguns casos, tanto as linhas quanto as colunas representam níveis de variáveis resposta. Tais casos caracterizam estudos onde o interesse é a associação entre variáveis. Representaremos a seguir uma tabela de contingência resultante da classificação de uma amostra de n indivíduos (plantas, animais, pessoas etc) em s subpopulações e r categorias de resposta (Tabela 3.1).

O valor Y_{ij} em cada casela é o número de indivíduos com a resposta j na amostra correspondente à subpopulação i . Outra maneira de representar os dados é através de

	1	2	...	j	...	r	
1	Y_{11}	Y_{12}	...	Y_{1j}	...	Y_{1r}	n_1
2	Y_{21}	Y_{22}	...	Y_{2j}	...	Y_{2r}	n_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
i	Y_{i1}	Y_{i2}	...	Y_{ij}	...	Y_{ir}	n_i
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
s	Y_{s1}	Y_{s2}	...	Y_{sj}	...	Y_{sr}	n_s

Tabela 3.1: Tabela de contingencia $s \times r$.

um vetor ($rs \times 1$) de observações

$$\mathbf{Y}_{rs \times 1} = (Y_{11}, \dots, Y_{1r}, \dots, Y_{s1}, \dots, Y_{sr}).$$

Segundo Koch *et al* (1984), para aplicação do método de quadrados mínimos ponderados é necessário dispor de um conjunto de observações arranjadas como um vetor ($u \times 1$), de funções $\mathbf{F} = (F_1(\mathbf{Y}), F_2(\mathbf{Y}), \dots, F_u(\mathbf{Y}))$ e um estimador consistente da matriz de covariâncias assintótica de \mathbf{F} . Além disso, a função \mathbf{F} deve ser definida de modo que sua matriz de covariâncias $\mathbf{V}_{\mathbf{F}}$ seja não singular.

Esse método é usualmente empregado na análise de dados categóricos modelados pelas distribuições de *produto de multinomiais* ou *Poisson*. Os estimadores dos parâmetros de interesse nessas distribuições têm propriedades assintóticas necessárias à aplicação do método delta (Apêndice B), usado para cálculo da distribuição assintótica de \mathbf{F} . Exemplos de dados modelados por essas distribuições serão abordados no capítulo 5.

3.1 Funções de resposta.

Seja o conjunto de observações $\mathbf{Y} \in IR^{rs}$ e $\hat{\theta}_n(\mathbf{Y}) \in IR^p$ um estimador do vetor de parâmetros que indexa a distribuição de \mathbf{Y} . Uma *função de resposta* é uma função definida como

$$F : IR^{rs} \mapsto IR^u$$

ou seja,

$$\mathbf{F}(\mathbf{Y}) = \begin{pmatrix} F_1(\mathbf{Y}) \\ F_2(\mathbf{Y}) \\ \vdots \\ F_u(\mathbf{Y}) \end{pmatrix}.$$

Quando há mais de uma categoria de resposta, o número de componentes de $\mathbf{F}(\mathbf{Y})$ não deve exceder o número de componentes de \mathbf{Y} linearmente independentes. Por exemplo, quando os dados são modelados pela distribuição produto de multinomiais, $u \leq (r-1)s$.

Usualmente, $\mathbf{F}(\cdot)$ é expressa indiretamente como função de \mathbf{Y} , através de $\hat{\theta}_n(\mathbf{Y})$, isto é, $\mathbf{F}(\cdot)$ é definida como uma função

$$F : IR^p \mapsto IR^u$$

que associa a cada vetor $\hat{\theta}_n(\mathbf{Y})$ em IR^p , o vetor

$$\mathbf{F}(\hat{\theta}_n(\mathbf{Y})) = \begin{pmatrix} F_1(\hat{\theta}_n(\mathbf{Y})) \\ F_2(\hat{\theta}_n(\mathbf{Y})) \\ \vdots \\ F_u(\hat{\theta}_n(\mathbf{Y})) \end{pmatrix} \equiv \begin{pmatrix} F_1(\hat{\theta}_n) \\ F_2(\hat{\theta}_n) \\ \vdots \\ F_u(\hat{\theta}_n) \end{pmatrix}.$$

As funções de resposta são definidas com base nos objetivos do investigador, mas, além de representar uma variável resposta de interesse, as funções devem possuir algumas características matemáticas adequadas à aplicação do método GSK. Essas características serão discutidas nas seções subseqüentes.

3.1.1 Escolha da função de resposta.

De um modo geral, a escolha da função de resposta baseia-se em um ou mais dos critérios abaixo:

a) *Objetivo da pesquisa.* As funções devem ser uma representação matemática da resposta cuja variabilidade busca-se explicar. Em muitos casos, as funções de resposta são definidas com base nas hipóteses que se deseja testar. Por exemplo, quando o interesse do estudo é testar hipóteses de associação multiplicativa, usam-se funções envolvendo transformações logarítmicas.

b) *Facilidade computacional na estimação de parâmetros:* quando o vetor de parâmetros pertence a um espaço paramétrico restrito, como é o caso do vetor π da distribuição produto de binomiais, o cálculo dos estimadores torna-se mais simples quando usam-se funções (transformações) de modo que $F_i(\pi) \in (-\infty, \infty)$ para todo $i = 1, 2, \dots, u$. Um exemplo desse tipo de transformação é a função logito, onde $F_i(\pi) = \ln\left(\frac{\pi_i}{1-\pi_i}\right)$:

c) *Busca de melhor ajuste para o modelo linear:* muitas vezes, pela natureza do fenômeno estudado, não se obtém bom ajuste de modelos lineares, quando usa-se os dados originais (não transformados). Em alguns casos, técnicas gráficas podem auxiliar na busca de uma transformação adequada.

Nos casos b e c, é importante verificar se a transformação resulta numa variável interpretável.

A função de interesse, cuja esperança assintótica deseja-se modelar, é usualmente expressa como função de um estimador do vetor de parâmetros (θ) que indexa a distribuição probabilística dos dados, denotado por $\hat{\theta}_n$.

O método geralmente utilizado para obtenção da matriz de covariâncias assintótica de $\mathbf{F}(\hat{\theta}_n)$, necessária para aplicação do método GSK, é a expansão de $\mathbf{F}(\hat{\theta}_n)$ em série de Taylor linear em torno de θ ,

$$\mathbf{F}(\hat{\theta}_n) = \mathbf{F}(\theta) + \mathbf{H}(\theta)(\hat{\theta}_n - \theta) + R(\hat{\theta}_n) \quad (3.1)$$

A notação $R(\hat{\theta}_n)$ refere-se ao resto da expansão e \mathbf{H} é uma matriz $u \times t$ cujo elemento (i,j) é a derivada da componente F_i de \mathbf{F} com respeito á j -ésima coordenada do vetor $\hat{\theta}_n = (\hat{\theta}_1^{(n)}, \hat{\theta}_2^{(n)}, \dots, \hat{\theta}_j^{(n)}, \dots, \hat{\theta}_p^{(n)})$.

Para uso da expansão definida em (3.1) é necessário que $\mathbf{F}(\cdot)$ tenha derivadas parciais de primeira ordem numa região contendo θ . A exigência sobre a existência de derivadas parciais contínuas até ordem dois, citada em Koch *et al* (1984), está relacionada com consistência do estimador da matriz de covariâncias assintótica de $\mathbf{F}(\hat{\theta}_n)$.

Assim, quando $\frac{\partial \mathbf{F}}{\partial \hat{\theta}_n}$ existe, $\hat{\theta}_n$ converge em probabilidade a θ e $R(\hat{\theta}_n)$ é de ordem de magnitude menor que $\|(\hat{\theta}_n - \theta)\|$, a função pode ser aproximada pela expansão até o termo de segunda ordem, isto é,

$$\mathbf{F}(\hat{\theta}_n) \simeq \mathbf{F}(\theta) + \mathbf{H}(\theta)(\hat{\theta}_n - \theta), \quad (3.2)$$

quando os tamanhos de amostra utilizados no cálculo das componentes de $\mathbf{F}(\hat{\theta}_n)$ são grandes.

Quando $\hat{\theta}_n$ é assintoticamente normal e (3.2) vale, o método delta multivariado pode ser utilizado para o cálculo da distribuição assintótica, e conseqüentemente, da esperança e covariância assintótica de $\mathbf{F}(\hat{\theta}_n)$. Isso ocorre quando os dados são modelados pelas distribuições *produto de Poisson* ou *multinomial*.

Segundo Koch *et al* (1984), o método de *quadrados mínimos ponderados* pode ser utilizado quando a função de interesse satisfaz:

- i) $\mathbf{F}(\hat{\theta}_n)$ deve ter derivadas parciais contínuas até ordem dois, numa região contendo o parâmetro;
- ii) $\mathbf{F}(\hat{\theta}_n)$ deve ser definida de modo que \mathbf{V}_F seja não singular.

Observe que se a condição (i) é satisfeita e $\hat{\theta}_n$ é um estimador consistente de θ , dispomos de um estimador consistente de \mathbf{V}_F . A condição (ii) assegura sua não singularidade. Essas implicações serão discutidas com maiores detalhes na seção seguinte.

3.1.2 Distribuição assintótica de F .

Usualmente, a função F não é expressa diretamente como função do conjunto de dados categóricos Y , mas de alguma estatística $\hat{\theta}_n(Y)$, geralmente um estimador consistente e não viciado do vetor de parâmetros que indexa a distribuição de Y . Uma técnica importante para calcular a distribuição assintótica de $F(\hat{\theta}_n)$ e conseqüentemente sua esperança e matriz de covariâncias assintóticas é o método delta (δ) multivariado descrito em Bishop *et al* (1977).

Seja $\hat{\theta}_n = (\hat{\theta}_1^{(n)}, \hat{\theta}_2^{(n)}, \dots, \hat{\theta}_p^{(n)})$ um vetor aleatório p -dimensional e $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ um vetor de parâmetros também p -dimensional. Consideremos que $\hat{\theta}_n$ e θ são tais que:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N_p(0, \Sigma(\theta)) \quad (3.3)$$

onde $\Sigma(\theta)$ é a matriz de covariância assintótica de $\hat{\theta}_n$ e n é o tamanho de amostra utilizado para o cálculo de cada componente de $\hat{\theta}_n$. No capítulo 5 será considerado o caso em que diferentes tamanhos de amostra são utilizados para o cálculo dos $\hat{\theta}_j$.

A expressão (3.3) equivale a afirmação de que, para n grande, $\hat{\theta}_n$ tem distribuição aproximada normal p -variada com média θ e matriz de covariâncias $n^{-1} \Sigma(\theta)$.

Agora suponha F uma função definida num subconjunto aberto do espaço p -dimensional tomando valores no espaço u -dimensional. Consideramos que F é diferenciável em θ , isto é, F tem a seguinte expansão quando $\hat{\theta}_n$ converge em probabilidade a θ ,

$$F_i(\hat{\theta}_n) = F_i(\theta) + \sum_{j=1}^p (\hat{\theta}_j^{(n)} - \theta_j) \left[\frac{\partial F_i}{\partial \hat{\theta}_j^{(n)}} \right]_{x=\theta} + o(\|\hat{\theta}_j^{(n)} - \theta_j\|) \quad (3.4)$$

para $i = 1, 2, \dots, u$.

Seja $H(\theta) = \left[\frac{\partial f}{\partial \theta} \right]$ uma matriz $u \times p$ cujo elemento (i,j) é a derivada parcial da i -ésima componente de $F = (F_1, F_2, \dots, F_u)$ com respeito à j -ésima coordenada de $\hat{\theta}_n = (\hat{\theta}_1^{(n)}, \hat{\theta}_2^{(n)}, \dots, \hat{\theta}_p^{(n)})$, avaliada em $\hat{\theta}_n = \theta$, isto é:

$$\mathbf{H}(\theta) = \left(\frac{\partial f}{\partial \theta} \right)_{ij} = \left[\frac{\partial f_i}{\partial \hat{\theta}_j^{(n)}} \right]_{\hat{\theta}_n = \theta} \quad (3.5)$$

então a expressão em (3.4) pode ser resumida como

$$\mathbf{F}(\hat{\theta}_n) = \mathbf{F}(\theta) + \mathbf{H}(\theta)(\hat{\theta}_n - \theta) + o(\|\hat{\theta}_n - \theta\|) \quad (3.6)$$

quando $\hat{\theta}_n$ converge em probabilidade a θ . A notação $o(\|\hat{\theta}_n - \theta\|)$ significa que o resto da expansão definida em (3.6) tem ordem de magnitude menor que $\|\hat{\theta}_n - \theta\|$.

Supondo que (3.3) e (3.6) valem, pelo teorema que resume o método delta multivariado (Teorema B5) a distribuição assintótica de $\mathbf{F}(\hat{\theta}_n)$ é dada por

$$\sqrt{n}[\mathbf{F}(\hat{\theta}_n) - \mathbf{F}(\theta)] \xrightarrow{D} N_u(\mathbf{0}, \mathbf{H}(\theta)\Sigma(\theta)\mathbf{H}'(\theta))$$

ou seja, para n grande, $\mathbf{F}(\hat{\theta}_n)$ tem distribuição aproximada normal multivariada com

$$\mathbf{E}(\mathbf{F}(\hat{\theta}_n)) \cong \mathbf{F}(\theta) \quad (3.7)$$

e matriz de covariâncias

$$\text{Var}(\mathbf{F}(\hat{\theta}_n)) \cong n^{-1}(\mathbf{H}(\theta)\Sigma(\theta)\mathbf{H}'(\theta)) \quad (3.8)$$

Quando $\mathbf{F}(\cdot)$ é uma transformação linear, isto é,

$$\mathbf{F}(\hat{\theta}_n) = \mathbf{A}\hat{\theta}_n$$

para alguma matriz \mathbf{A} , $u \times p$ de posto linha completo, a esperança e covariância de $\mathbf{F}(\hat{\theta}_n)$ calculadas pelo método delta multivariado são exatas. Apenas a distribuição é aproximada, pois:

$$\mathbf{E}(\mathbf{F}(\hat{\theta}_n)) = \mathbf{E}(\mathbf{A}\hat{\theta}_n) = \mathbf{A}\mathbf{E}(\hat{\theta}_n) = \mathbf{A}\theta = \mathbf{F}(\theta)$$

e

$$\text{Var}(\mathbf{F}(\hat{\theta}_n)) = \text{Var}(\mathbf{A}\hat{\theta}_n) = \mathbf{A}\text{Var}(\hat{\theta}_n)\mathbf{A}' = n^{-1}(\mathbf{H}(\theta)\Sigma(\theta)\mathbf{H}'(\theta))$$

Substituindo θ por $\hat{\theta}_n$ em (3.7) e (3.8) obtemos estimadores da esperança e matriz de covariâncias assintótica de $(\mathbf{F}(\hat{\theta}_n))$. Denotaremos esses estimadores por \mathbf{F} e \mathbf{V}_F , respectivamente. Quando $\hat{\theta}_n$ é um estimador consistente de θ e as componentes dos estimadores da esperança assintótica de \mathbf{F} (denotada por $\mathbf{E}_A(\mathbf{F})$) e \mathbf{V}_F são funções contínuas de $\hat{\theta}_n$, estes também são estimadores consistentes de \mathbf{F} e da variância assintótica de \mathbf{F} , respectivamente. Se $\mathbf{F}(\cdot)$ possui derivadas parciais contínuas até ordem dois, as componentes de $\mathbf{F}(\cdot)$ e $\mathbf{H}(\cdot)$ são contínuas e está assegurada a consistência de \mathbf{V}_F .

A não singularidade de \mathbf{V}_F depende apenas do posto-linha de \mathbf{H} . Se $\mathbf{F}(\cdot)$ é tal que as u linhas de \mathbf{H} são linearmente independentes,

$$\mathbf{V}_F = n^{-1}(\mathbf{H}(\hat{\theta}_n)\Sigma(\hat{\theta}_n)\mathbf{H}'(\hat{\theta}_n))$$

é não singular (tem posto u).

3.2 O modelo linear geral

O modelo linear (geral) usualmente postulado para descrever as funções compreendidas em \mathbf{F} , é um modelo de coordenadas fixas da forma:

$$E_A(\mathbf{F}) = \mathbf{X}\beta \tag{3.9}$$

onde \mathbf{X} é uma matriz ($u \times t$) de constantes conhecidas de posto $t \leq u$ (matriz do modelo) e β é um vetor ($t \times 1$) de parâmetros desconhecidos.

Esse modelo também pode ser expresso postulando-se que $E_A(\mathbf{F})$ pertence a um subespaço $V \subset R^n$ gerado pelas colunas da matriz \mathbf{X} (ou por outra matriz \mathbf{B} qualquer tal que $C(\mathbf{B}) = C(\mathbf{X})$). Essa é a formulação livre de coordenadas e também pode ser

expressa através de um conjunto de restrições lineares sobre $E_A(\mathbf{F})$.

Seja \mathbf{W} uma matriz $(u-t) \times u$ de posto $u-t$ cujo espaço-linha é ortogonal ao subespaço V e \mathbf{X} uma base de V . Então, o espaço linha de \mathbf{W} é ortogonal ao espaço-coluna de \mathbf{X} , isto é $\mathbf{WX} = \mathbf{0}$. Dessa forma, se (3.9) vale,

$$\mathbf{W}[E_A(\mathbf{F})] = \mathbf{WX}\beta = \mathbf{0}$$

para todo $\beta \in \Theta$, e o modelo (3.9) pode ser postulado alternativamente como:

$$\mathbf{W}[E_A(\mathbf{F})] = \mathbf{0} \tag{3.10}$$

3.3 Ajuste do modelo

Após o cálculo dos valores das componentes do vetor \mathbf{F} a partir dos dados observados é necessário verificar a qualidade do ajuste do modelo.

No método de quadrados mínimos ponderados, a qualidade do ajuste é medida através do χ^2 de Wald. Quando os dados têm distribuição de Poisson ou multinomial essa estatística é assintoticamente equivalente à estatística da razão da verossimilhança, no sentido de que têm a mesma distribuição aproximada para tamanhos de amostra grandes em cada subpopulação, que é χ^2 com $(u-t)$ graus de liberdade.

A estatística de Wald é dada pela expressão:

$$Q_W = (\mathbf{WF})'(\mathbf{WV}_F\mathbf{W}')^{-1}(\mathbf{WF})$$

Considerando que o modelo (3.9) é válido, espera-se que \mathbf{WF} resulte num vetor próximo do vetor nulo. A medida de distância é usada para indicar se $\mathbf{W}[E_A(\mathbf{F})]$ está próximo ou distante do vetor nulo é a norma V_C^{-1} de seu estimador consistente \mathbf{WF} onde $V_C = \mathbf{WV}_F\mathbf{W}'$ é a estimativa da matriz de covariâncias de \mathbf{WF} . O uso dessa norma incorpora ao cálculo de Q_W a estrutura de covariâncias de \mathbf{WF} , dando maior peso às componentes de menor variância.

Uma expressão alternativa para a estatística de Wald é:

$$Q_W = (\mathbf{F} - \mathbf{P}_V \mathbf{F})' \mathbf{V}_F^{-1} (\mathbf{F} - \mathbf{P}_V \mathbf{F}) \quad (3.11)$$

onde $\mathbf{P}_V \mathbf{F}$ é a projeção oblíqua de $\mathbf{F} \in \mathbb{R}^n$ no subespaço V tal que a norma- \mathbf{V}_F^{-1} do vetor $(\mathbf{F} - \mathbf{P}_V \mathbf{F})$ é mínima.

3.4 Análise de resíduos

No modelo linear (comum) a análise de resíduos busca evidências sobre o possível não cumprimento das presunções de normalidade e homocedasticidade dos erros e necessidade de inclusão de novas variáveis preditoras. No modelo linear geral, a estrutura de covariâncias não pressupõe homocedasticidade nem independência dos erros. Nesse caso, a análise de resíduos investiga a distribuição assintótica dos erros, e busca possíveis associações dos resíduos do modelo definido em (3.9) com variáveis não incluídas no modelo.

O modelo (3.9) pode ser escrito de forma alternativa como:

$$\mathbf{F} = \mathbf{X}\beta + \epsilon. \quad (3.12)$$

O vetor

$$\epsilon = \mathbf{F} - \mathbf{X}\beta \quad (3.13)$$

é chamado vetor de erros e representa os desvios entre as componentes do vetor de estatísticas \mathbf{F} e sua esperança assintótica sob o modelo. O vetor de erros é estimado pelo vetor de resíduos $\mathbf{e} = \mathbf{F} - \mathbf{X}\hat{\beta}$ onde $\hat{\beta}$ é um estimador qualquer de β .

De (3.13), observa-se que ϵ é uma transformação linear de \mathbf{F} . Então, quando \mathbf{F} tem distribuição assintoticamente normal, e pelo Teorema B6, a distribuição de ϵ é também assintoticamente normal, com

$$\mathbf{E}_A(\epsilon) = \mathbf{E}_A(\mathbf{F} - \mathbf{X}\beta) = \mathbf{E}_A(\mathbf{F}) - \mathbf{X}\beta = \mathbf{0}, \quad (3.14)$$

se o modelo (3.9) vale. Note que

$$\text{Var}_A(\mathbf{F} - \mathbf{X}\beta) = \text{Var}_A \mathbf{F}.$$

O gráfico dos resíduos em papel probabilístico pode ser útil para verificar a normalidade dos erros. Quando o número de componentes da função de resposta é pequeno (como é o caso dos exemplos do capítulo 5), o uso desse método é limitado. O pequeno número de pontos no gráfico de resíduos dificilmente fornecerá evidência suficiente contra a hipótese de normalidade.

Para testar se novas variáveis devem ser incluídas no modelo, colunas adicionais correspondentes a essas variáveis são acrescentadas à matriz \mathbf{X} . Essas colunas devem ser tais que a matriz aumentada $[\mathbf{X}, \mathbf{X}_E]$ tenha posto coluna completo.

Busca-se então investigar se variáveis candidatas têm associação com a variabilidade não explicada pelo conjunto de preditoras do modelo anterior (resíduos). Isso corresponde à hipótese de que o vetor de erros pode ser expresso como combinação linear das colunas de \mathbf{X}_E , ou seja

$$\epsilon \in C(\mathbf{X}_E) \quad (3.15)$$

onde \mathbf{X}_E é uma matriz $u \times t_E$ tal que o posto de $(\mathbf{X}, \mathbf{X}_E)$ é $t + t_E$.

Para testar essa hipótese usa-se a estatística :

$$Q_{X_E} = \mathbf{e}'\mathbf{X}_E(\mathbf{X}_E'\mathbf{V}_e\mathbf{X}_E)^{-1}\mathbf{X}_E'\mathbf{e} = (\mathbf{F} - \hat{\mathbf{F}})'\mathbf{X}_E(\mathbf{X}_E'\mathbf{V}_{(\mathbf{F}-\hat{\mathbf{F}})}\mathbf{X}_E)^{-1}\mathbf{X}_E'(\mathbf{F} - \hat{\mathbf{F}}),$$

onde

$$\mathbf{V}_e = \mathbf{V}_{(\mathbf{F}-\hat{\mathbf{F}})} = \mathbf{V}_F - \mathbf{V}_{\hat{\mathbf{F}}} = \mathbf{V}_F - \mathbf{X}\mathbf{V}_b\mathbf{X}'.$$

\mathbf{V}_b é um estimador da matriz de covariâncias assintótica do estimador de quadrados mínimos ponderados de β e será definido na seção 2.6.

A estatística Q_{X_E} corresponde à distância entre o vetor de resíduos do modelo (3.9),

$$\mathbf{e} = \mathbf{F} - \hat{\mathbf{F}}$$

e sua projeção no espaço coluna de \mathbf{X}_E

$$\mathbf{P}_{\mathbf{X}_E}(\mathbf{e}) = \mathbf{X}_E(\mathbf{X}'_E\mathbf{V}_e\mathbf{X}_E)^{-1}\mathbf{X}'_E\mathbf{V}_e^{-1}\mathbf{e}$$

ponderada pela matriz de covariâncias assintótica de \mathbf{e} .

Sob a hipótese definida em (3.15), Q_{X_E} tem distribuição assintótica χ^2 com t_E graus de liberdade (Teorema B8).

3.5 Método de quadrados mínimos ponderados.

A escolha do estimador $\hat{\beta}$ de β no modelo linear geral pode basear-se em diferentes critérios, entre eles o critério da *máxima verossimilhança* e o de *quadrados mínimos ponderados* (QMP). Quando os dados analisados têm distribuição de Poisson ou multinomial, os estimadores de QMP são assintoticamente equivalentes aos estimadores de máxima verossimilhança, no sentido de que possuem a mesma distribuição aproximada quando os tamanhos das amostras utilizadas para o cálculo de \mathbf{F} são suficientemente grandes.

O critério de Q.M.P. baseia-se na escolha de β que minimize a soma de quadrados de resíduos ponderada pelo inversa do estimador da matriz de covariâncias assintóticas :

$$Q(\beta) = (\mathbf{F} - \mathbf{X}\beta)' \mathbf{V}_F^{-1} (\mathbf{F} - \mathbf{X}\beta) \quad (3.16)$$

O estimador de β baseado nesse critério será denotado por \mathbf{b} e chamado *estimador de quadrados mínimos ponderados* de β .

3.5.1 Estimação do vetor de parâmetros do modelo.

Uma das formas de encontrar \mathbf{b} é diferenciar $Q(\beta)$ em relação a β em $\beta = \mathbf{b}$ e igualar a expressão resultante a zero:

$$\frac{\partial Q(\beta)}{\partial \beta} = (\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})\mathbf{b} - \mathbf{X}'\mathbf{V}_F^{-1}\mathbf{F} = 0 \quad (3.17)$$

Daí resulta o sistema conhecido como *equações de Aitken*:

$$(\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{V}_F^{-1}\mathbf{F} \quad (3.18)$$

Como \mathbf{X} é uma matriz $n \times p$ com posto-coluna completo p e \mathbf{V}_F é $u \times u$ não singular, $\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X}$ é também não singular e a solução do sistema (3.18) é :

$$\mathbf{b} = (\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{F}. \quad (3.19)$$

Dessa forma o valor mínimo de $Q(\beta)$ ocorre quando $\beta = \mathbf{b}$, isto é, a expressão

$$\min_{\beta \in \mathbb{R}^p} Q(\beta) = (\mathbf{F} - \mathbf{X}\mathbf{b})\mathbf{V}_F^{-1}(\mathbf{F} - \mathbf{X}\mathbf{b})$$

corresponde à estatística de Wald definida em (3.11).

Quando $\mathbf{V}_F = \sigma^2\mathbf{I}$ e $\mathbf{F} = \mathbf{y}$ temos o caso descrito no capítulo anterior onde \mathbf{b} é escolhido de modo que a norma L_2 do vetor de resíduos seja minimizada. Essa norma corresponde à distância euclidiana entre o vetor de observações \mathbf{y} e sua projeção ortogonal no $C(\mathbf{X})$.

No critério de Q.M.P. busca-se o valor de \mathbf{b} tal que a distância entre o vetor \mathbf{F} e sua projeção no $C(\mathbf{X})$, $\hat{\mathbf{F}} = \mathbf{X}\mathbf{b}$ tenha mínima norma $-\mathbf{V}_F^{-1}$. O uso dessa norma permite uma ponderação das componentes do vetor de resíduos que considera a estrutura de

covariâncias de \mathbf{F} .

3.5.2 Estimação da função de resposta.

De (3.19), o vetor de estimativas $\hat{\mathbf{F}} = \mathbf{X}\mathbf{b}$ pode ser escrito como:

$$\hat{\mathbf{F}} = \mathbf{X}(\mathbf{X}'\mathbf{V}_{\mathbf{F}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_{\mathbf{F}}^{-1}\mathbf{F} \quad (3.20)$$

Na postulação de coordenadas livres do modelo (3.9) afirma-se que $E_A(\mathbf{F})$ pertence ao subespaço V (gerado pelas colunas de \mathbf{X} ou outra base qualquer). A matriz

$$\mathbf{P}_V = \mathbf{X}(\mathbf{X}'\mathbf{V}_{\mathbf{F}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_{\mathbf{F}}^{-1}$$

corresponde à matriz do projetor \mathbf{P}_V que aplicado em $\mathbf{F} \in R^u$ projeta-o no subespaço V segundo o ponto de tangência do elipsóide definido por $Q(\beta)$ em (3.16). Dessa forma, a função de resposta estimada pelo modelo pode ser expressa como

$$\hat{\mathbf{F}} = \mathbf{P}_V\mathbf{F}.$$

Observa-se que \mathbf{P}_V é idempotente (não simétrica) e portanto, pelo Teorema A18, \mathbf{P}_V é um projetor não ortogonal, exceto quando $\mathbf{V}_{\mathbf{F}} = \mathbf{I}$.

Quando o modelo proposto explica adequadamente a variabilidade entre as componentes de \mathbf{F} , o próximo passo é testar hipóteses sobre os parâmetros. supondo o modelo correto.

3.6 Propriedades dos estimadores de quadrados mínimos ponderados.

Seja o modelo linear geral

$$\mathbf{E}_A \mathbf{F}(\hat{\theta}_n) = \mathbf{X}\beta \quad (3.21)$$

onde $\hat{\theta}_n$ é um estimador consistente do parâmetro (θ) que indexa a distribuição dos dados e \mathbf{X} e β são conforme definidos em (3.9).

Se $\hat{\theta}_n$ é uma estatística tal que a condição (3.3) é satisfeita e $\mathbf{F}(\hat{\theta}_n) \equiv \mathbf{F}$ pode ser escrita segundo a expansão definida em (3.2), pelo método delta multivariado,

$$\mathbf{F}(\hat{\theta}_n) \stackrel{a}{\approx} \mathbf{N}_u(\mathbf{F}(\theta), n^{-1} \mathbf{H}(\theta) \boldsymbol{\Sigma}(\theta) \mathbf{H}'(\theta)). \quad (3.22)$$

De (3.19), o estimador de quadrados mínimos ponderados é dado pela expressão:

$$\mathbf{b} = (\mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}_F^{-1} \mathbf{F}. \quad (3.23)$$

Então, \mathbf{b} pode ser escrito como uma transformação linear de \mathbf{F} , e pelo Teorema B6, \mathbf{b} tem distribuição assintótica normal multivariada .

De (3.23), a esperança assintótica de \mathbf{b} é:

$$\mathbf{E}_A(\mathbf{b}) = \mathbf{E}_A[(\mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}_F^{-1} \mathbf{F}] = (\mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}_F^{-1} \mathbf{E}_A(\mathbf{F})$$

Supondo o modelo (3.9) verdadeiro,

$$\mathbf{E}_A(\mathbf{b}) = (\mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X} \beta = \beta \quad (3.24)$$

A matriz de covariâncias assintótica de \mathbf{b} é dada pela expressão

$$\begin{aligned} \text{Var}_A(\mathbf{b}) &= \text{Var}_A[(\mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}_F^{-1} \mathbf{F}] = \\ &= [(\mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}_F^{-1}] \text{Var}_A(\mathbf{F}) [(\mathbf{X}' \mathbf{V}_F^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}_F^{-1}]' \end{aligned} \quad (3.25)$$

Substituindo $\text{Var}_A(\mathbf{F})$ em (3.25) por seu estimador consistente \mathbf{V}_F obtém-se um

estimador consistente de $\text{Var}_A(\mathbf{b})$, denotado por \mathbf{V}_b :

$$\begin{aligned}
 \mathbf{V}_b &= [(\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_F^{-1}]\mathbf{V}_F[(\mathbf{V}_F)^{-1}]\mathbf{X}[(\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}]' = \\
 &= (\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}\mathbf{X}'[(\mathbf{V}_F)^{-1}]\mathbf{X}[(\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}]' = \\
 &= (\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}(\mathbf{X}'(\mathbf{V}_F)^{-1}\mathbf{X})'[(\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}]' = \\
 &\qquad\qquad\qquad (\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1} \qquad\qquad\qquad (3.26)
 \end{aligned}$$

Se $\mathbf{F}(\cdot)$ é uma transformação linear, a esperança e a matriz de covariâncias de \mathbf{b} são exatas, apenas sua distribuição é aproximada.

Quando os dados analisados têm distribuição Poisson ou multinomial, o método de estimação quadrado mínimos ponderados produz estimadores BAN (Best asymptotically normal), isto é, estimadores assintoticamente normais e assintoticamente eficientes.

3.7 Testes de hipóteses lineares

Hipóteses sobre um vetor de parâmetros $\beta \in \Theta$, são ditas *hipóteses lineares* quando podem ser expressas através de um conjunto de restrições lineares sobre o espaço paramétrico Θ .

Essas hipóteses podem ser formuladas diretamente sobre a $\mathbf{E}_A(\mathbf{F})$ sem necessidade de postulação anterior de um modelo. Nesse caso, a hipótese é descrita por

$$\mathbf{W}\mathbf{E}_A(\mathbf{F}) = \mathbf{0}$$

onde \mathbf{W} é uma matriz $c \times u$ de posto c .

Usualmente, as hipóteses são formuladas após o ajuste de um modelo, quando este

é considerado adequado. Assim, as hipóteses representam restrições adicionais sobre o vetor de parâmetros, além das postuladas pelo modelo.

Seja $C\beta = 0$ um conjunto de restrições lineares sobre Θ que os parâmetros β devem satisfazer, onde C é uma matriz $c \times t$. Para que não haja restrições redundantes, as linhas da matriz C devem ser linearmente independentes e conseqüentemente, a matriz C deve ter posto-linha completo, isto é $\text{posto}(C) = c$.

Sob a hipótese $H_0 : C\beta = 0$ espera-se que a estimativa Cb do contraste $C\beta$ esteja próxima do vetor nulo. Para testar essa hipótese pode ser usada a estatística de Wald (Wald (1953)):

$$Q_C = b' C' (V_c)^{-1} C b \quad (3.27)$$

que quantifica essa proximidade através da norma $-(V_C)^{-1}$ do vetor Cb . V_C é a estimativa da matriz de covariâncias de Cb , denotada por $\text{Var}(Cb)$ e expressa por:

$$\text{Var}(Cb) = C \text{Var}(b) C' \quad (3.28)$$

onde $\text{Var}(b)$ é a matriz de covariâncias do estimador de quadrados mínimos ponderados b .

De (3.27) e (3.28) observa-se que a exigência de que a matriz C tenha posto linha completo, é necessária para que V_c seja inversível, e conseqüentemente para o cálculo de Q_C .

De (3.26), um estimador consistente de $\text{Var}(b)$ é:

$$V_b = (X' V_F^{-1} X)^{-1} \quad (3.29)$$

onde V_F^{-1} é um estimador consistente da matriz de covariâncias assintótica de F . V_b é uma matriz $p \times p$ de posto completo e C tem posto linha completo, o que garante a não singularidade de $\text{Var}(Cb)$, exigida para o cálculo de Q_C .

Dessa forma a expressão (3.27) resulta em:

$$\begin{aligned} Q_C &= \mathbf{b}'\mathbf{C}'(\mathbf{C}\mathbf{V}_b\mathbf{C}')^{-1}\mathbf{C}\mathbf{b} & (3.30) \\ &= \mathbf{b}'\mathbf{C}'[\mathbf{C}(\mathbf{X}'\mathbf{V}_F^{-1}\mathbf{X})^{-1}\mathbf{C}']^{-1}\mathbf{C}\mathbf{b} \end{aligned}$$

Pelo Teorema B8, a estatística Q_C tem distribuição assintótica χ^2 com c graus de liberdade. Valores altos de Q_C fornecem evidência para rejeição de H_0 . Para decidir se um valor é alto ou baixo calcula-se o respectivo valor p com base na distribuição assintótica de Q_C .

Capítulo 4

Modelagem probabilística das tabelas de contingência.

As principais distribuições utilizadas na modelagem probabilística de tabelas de contingência são a *distribuição multinomial* e a *distribuição de Poisson*. Os parâmetros que indexam essas distribuições possuem estimadores consistentes, não viciados e assintoticamente normais. Essas propriedades assintóticas propiciam a utilização do método delta no cálculo da distribuição assintótica de funções particulares desses estimadores.

Neste capítulo caracterizam-se situações em que os dados objeto de análise podem ser adequadamente modelados pelas distribuições multinomial ou Poisson. São derivadas distribuições assintóticas dos estimadores dos parâmetros que indexam tais distribuições e obtidos estimadores consistentes das matrizes de covariâncias assintóticas de funções desses estimadores. Análises detalhadas de exemplos dessas situações serão apresentadas no capítulo 5.

4.1 Dados modelados pela distribuição produto de multinomiais,

Considere o conjunto de dados categóricos apresentado na Tabela 3.1. Nessa tabela, as linhas correspondem a subpopulações e as colunas a categorias de resposta. Os totais marginais n_1, n_2, \dots, n_s constituem os tamanhos de amostra em cada subpopulação e as

variáveis aleatórias Y_{ij} em cada casela representam o número de indivíduos na amostra correspondente à subpopulação i que apresentaram a resposta j . Essas amostras são conceitualmente representativas de subpopulações infinitas e as tendências de cada indivíduo em apresentar a j -ésima resposta são consideradas mutuamente independentes. Além disso, a probabilidade de um indivíduo apresentar a resposta j é constante para indivíduos dentro da mesma subpopulação.

Considerando válidas as afirmações acima, o vetor aleatório $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{ir})$ tem distribuição multinomial com parâmetros n_i e $\pi_i = (\pi_{i1}, \dots, \pi_{ir})$ onde π_{ij} é a probabilidade de um indivíduo selecionado ao acaso na subpopulação i apresentar a categoria de resposta j . A função de probabilidade de \mathbf{Y}_i é

$$P(Y_{i1} = y_{i1}, \dots, Y_{ir} = y_{ir}) = (n_i)! \prod_{j=1}^r \frac{\pi_{ij}^{y_{ij}}}{(y_{ij})!} \quad (4.1)$$

com $\sum_{j=1}^r y_{ij} = n_i$ e $\sum_{j=1}^r \pi_{ij} = 1$, $\pi_{ij} \in (0, 1)$ para todo $i = 1, 2, \dots, s$ e $j = 1, 2, \dots, r$.

$\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_s$ são vetores aleatórios independentes com $\mathbf{Y}_i \sim \text{Multinomial}(n_i, \pi_i)$. Então o vetor $\mathbf{Y} = (Y_{11}, \dots, Y_{1r}, Y_{21}, \dots, Y_{2r}, \dots, Y_{s1}, \dots, Y_{sr})$ tem distribuição produto de multinomiais com função de probabilidade

$$P(Y_{11} = y_{11}, \dots, Y_{sr} = y_{sr}) = \prod_{i=1}^s (n_i)! \prod_{j=1}^r \frac{\pi_{ij}^{y_{ij}}}{(y_{ij})!} \quad (4.2)$$

O vetor de esperanças e a matriz de covariâncias de \mathbf{Y} são:

$$\mathbf{E}(\mathbf{Y}) = \mathbf{E} \begin{pmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{sr} \end{pmatrix} = \begin{pmatrix} n\pi_{11} \\ n\pi_{12} \\ \vdots \\ n\pi_{sr} \end{pmatrix}$$

e

$$\text{Var}(\mathbf{Y}) = \begin{bmatrix} V_1 & 0 & 0 & 0 \\ 0 & V_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & V_s \end{bmatrix} \quad (4.3)$$

onde

$$\mathbf{V}_i = n_i \begin{bmatrix} \pi_{i1}(1 - \pi_{i1}) & -\pi_{i1}\pi_{i2} & \cdots & -\pi_{i1}\pi_{ir} \\ -\pi_{i1}\pi_{i2} & \pi_{i2}(1 - \pi_{i2}) & \cdots & -\pi_{i2}\pi_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ -\pi_{i1}\pi_{ir} & -\pi_{i2}\pi_{ir} & \cdots & \pi_{ir}(1 - \pi_{ir}) \end{bmatrix}.$$

Um estimador consistente e não viciado de π é $\mathbf{p} = (p_{11}, p_{12}, \dots, p_{sr})$ onde $p_{ij} = Y_{ij}/n_i$. Substituindo π_{ij} por p_{ij} em (4.3), obtêm-se um estimado consistente de $\text{Var}(\mathbf{Y})$ que será denotado por \mathbf{V}_Y . O vetor de esperanças e a matriz de covariâncias de \mathbf{p} são:

$$\mathbf{E}(\mathbf{p}) = \mathbf{E} \begin{pmatrix} Y_{11}/n_1 \\ Y_{12}/n_1 \\ \vdots \\ Y_{sr}/n_s \end{pmatrix} = \begin{pmatrix} \pi_{11} \\ \pi_{12} \\ \vdots \\ \pi_{sr} \end{pmatrix} \quad (4.4)$$

$$\mathbf{V}_p(\pi) = \begin{bmatrix} W_1 & 0 & 0 & 0 \\ 0 & W_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & W_s \end{bmatrix} \quad (4.5)$$

onde

$$\mathbf{W}_i = \frac{1}{n_i} \begin{bmatrix} \pi_{i1}(1 - \pi_{i1}) & -\pi_{i1}\pi_{i2} & \cdots & -\pi_{i1}\pi_{ir} \\ -\pi_{i1}\pi_{i2} & \pi_{i2}(1 - \pi_{i2}) & \cdots & -\pi_{i2}\pi_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ -\pi_{i1}\pi_{ir} & -\pi_{i2}\pi_{ir} & \cdots & \pi_{ir}(1 - \pi_{ir}) \end{bmatrix}. \quad (4.6)$$

corresponde à matriz de covariâncias do vetor de proporções amostrais na subpopulação i , $\mathbf{p}_i = (p_{i1}, p_{i2}, \dots, p_{ir})$.

Denote $\mathbf{E}(\mathbf{p}_i) = \boldsymbol{\pi}_i$. Pelo teorema central do limite para vetores aleatórios (Teorema B3),

$$\sqrt{n_i}(\mathbf{p}_i - \boldsymbol{\pi}_i) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma}_i) \quad (4.7)$$

onde $\boldsymbol{\Sigma}_i$ é a matriz de covariâncias assintótica \mathbf{p}_i ,

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \pi_{i1}(1 - \pi_{i1}) & -\pi_{i1}\pi_{i2} & \cdots & -\pi_{i1}\pi_{ir} \\ -\pi_{i1}\pi_{i2} & \pi_{i2}(1 - \pi_{i2}) & \cdots & -\pi_{i2}\pi_{ir} \\ \vdots & \vdots & \ddots & \vdots \\ -\pi_{i1}\pi_{ir} & -\pi_{i2}\pi_{ir} & \cdots & \pi_{ir}(1 - \pi_{ir}) \end{bmatrix} \quad (4.8)$$

Seja $\mathbf{n} = (n_1, n_2, \dots, n_s)$ um vetor coluna cujas componentes são os tamanhos de amostra em cada subpopulação. Esse vetor pode ser reescrito como

$$\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_s \end{pmatrix} = \begin{pmatrix} n/k_1 \\ n/k_2 \\ \vdots \\ n/k_s \end{pmatrix}$$

onde $n = \sum_{i=1}^s n_i$ e k_1, k_2, \dots, k_s são constantes conhecidas. Sejam ainda,

$$\mathbf{D}_n = \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & D_s \end{bmatrix} \quad (4.9)$$

uma matriz bloco diagonal $sr \times sr$, onde \mathbf{D}_i é uma matriz diagonal $r \times r$.

$$\mathbf{D}_i = \begin{bmatrix} \frac{n}{k_i} & 0 & \cdots & 0 \\ 0 & \frac{n}{k_i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{n}{k_i} \end{bmatrix}.$$

Então o vetor \mathbf{p} , que resulta da concatenação vertical de $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_s$ satisfaz

$$(\mathbf{D}_n)^{1/2}(\mathbf{p} - \boldsymbol{\pi}) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma}_p(\boldsymbol{\pi})) \quad (4.10)$$

onde $\boldsymbol{\Sigma}_p(\boldsymbol{\pi})$ é uma matriz bloco diagonal $sr \times sr$,

$$\boldsymbol{\Sigma}_p(\boldsymbol{\pi}) = \begin{bmatrix} \Sigma_1 & 0 & 0 & 0 \\ 0 & \Sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \Sigma_s \end{bmatrix} \quad (4.11)$$

com Σ_i definida em (4.8).

Geralmente, quando os dados são modelados pela distribuição multinomial, é de interesse explicar a variabilidade entre componentes de funções do vetor de proporções amostrais \mathbf{p} . O modelo é usualmente postulado como

$$\mathbf{E}_A(\mathbf{F}(\mathbf{p})) = \mathbf{X}\boldsymbol{\beta}$$

onde \mathbf{X} é uma matriz $u \times p$ de constantes conhecidas com posto $(\mathbf{X}) = p$ e $\boldsymbol{\beta}$ é o vetor $p \times 1$ de parâmetros desconhecidos do modelo.

Satisfeita a condição definida em (4.10), se $\mathbf{F}(\mathbf{p})$ é tal que (3.2) vale, o método delta pode ser aplicado para o cálculo da esperança e covariância assintóticas de $\mathbf{F}(\mathbf{p})$. De (3.5) e (3.6) obtêm-se:

$$\mathbf{E}_A[\mathbf{F}(\mathbf{p})] = \mathbf{F}(\boldsymbol{\pi}) \quad (4.12)$$

$$\text{Var}_A[\mathbf{F}(\mathbf{p})] = (\mathbf{D}_n)^{-1} \mathbf{H}(\pi) [\Sigma_p(\pi)] \mathbf{H}'(\pi) \quad (4.13)$$

onde \mathbf{D}_n , $\mathbf{H}(\pi)$ e $\Sigma_p(\pi)$ estão definidos em (4.9), (3.5) e (4.11), respectivamente.

Substituindo \mathbf{p} por π em (4.13) obtêm-se um estimador consistente da matriz de covariâncias assintótica de $\mathbf{F}(\mathbf{p})$, que será denotado por :

$$\mathbf{V}_{\mathbf{F}(\mathbf{p})} = \mathbf{H}(\mathbf{p}) [\Sigma_p] \mathbf{H}'(\mathbf{p})$$

Para garantir que a matriz de covariâncias $\mathbf{V}_{\mathbf{F}(\mathbf{p})}$, obtida utilizando o método delta, é não singular, $\mathbf{F}(\mathbf{p})$ deve ser tal que $\mathbf{H}(\mathbf{p})$ tenha posto linha completo.

Satisfeita essa condição, dispõe-se de um estimador consistente de $\text{Var}_A[\mathbf{F}(\mathbf{p})]$, exigido para aplicação do método de *quadrados mínimos ponderados*.

4.2 Dados modelados pela distribuição produto de Poisson.

A distribuição de Poisson é freqüentemente utilizada para modelar situações onde a variável resposta de interesse é discreta, ordinal, assumindo valores inteiros positivos. É adequada para descrever dados relativos à contagem de eventos que ocorrem independente e aleatoriamente em alguma unidade de exposição fixa (tempo, área, volume, etc) com taxa de ocorrência constante.

James (1981) caracteriza de um modo formal um processo de Poisson através das seguintes hipóteses:

1. Hipótese 1. (*Incrementos estacionários*). A probabilidade da ocorrência de k eventos no intervalo $(s, s + t)$, depende somente de t e não de s ;
2. Hipótese 2. (*Incrementos independentes*). Os números de ocorrências em intervalos disjuntos da unidade de exposição são independentes;

SUBPOPULAÇÃO	RESPOSTA
1	Y_1
2	Y_2
\vdots	\vdots
i	Y_i
\vdots	\vdots
s	Y_s

Tabela 4.1: Número de ocorrências na subpopulação i para $i=1,2,\dots,s$.

3. Hipótese 3. Os eventos ocorrem sozinhos e não simultaneamente.

Haight (1967), *apud* por Plackett (1974) ressalta que a distribuição de Poisson é derivada a partir de princípios teóricos elementares, com um mínimo de pressuposições, em particular, pode ser obtida como limite das distribuições binomial ou hipergeométrica.

Suponha que Y_1, Y_2, \dots, Y_s (Tabela 4.2) são conceitualmente representativos de subpopulações infinitas $1, 2, \dots, s$, respectivamente no sentido de equivalência com uma amostra aleatória estratificada. Considere ainda que Y_1, Y_2, \dots, Y_s são mutuamente independentes com distribuição de Poisson (μ_i) para todo $i = 1, 2, \dots, s$.

Então o vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_s)$ tem distribuição produto de Poisson com vetor de parâmetros $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_s)$. A função de probabilidade de \mathbf{Y} é

$$P(Y_1 = y_1, Y_2 = y_2, \dots, Y_s = y_s) = \prod_{i=1}^s \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}$$

com $y_i \in \{0, 1, 2, \dots\}$ e $\mu_i \in (0, \infty)$ para todo $i = 1, 2, \dots, s$.

O vetor de esperanças e a matriz de covariâncias de \mathbf{Y} são:

$$\mathbf{E}(\mathbf{Y}) = \mathbf{E} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_s \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_s \end{pmatrix} \quad (4.14)$$

$$\mathbf{V}_Y(\boldsymbol{\mu}) = \mathbf{D}_\mu = \begin{bmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_s \end{bmatrix} \quad (4.15)$$

Seja $\mathbf{N} = (N_1, N_2, \dots, N_s)$ um vetor conhecido de *medidas de exposição* (tempo, tamanho da população sob risco, área geográfica, volume, etc). Definimos

$$\boldsymbol{\lambda} = \mathbf{D}_N^{-1} \boldsymbol{\mu} = \begin{bmatrix} \frac{1}{N_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{N_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{N_s} \end{bmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_s \end{pmatrix} = \begin{pmatrix} \frac{\mu_1}{N_1} \\ \frac{\mu_2}{N_2} \\ \vdots \\ \frac{\mu_s}{N_s} \end{pmatrix} \quad (4.16)$$

o vetor de parâmetros de razão onde $\lambda_i = \mu_i/N_i$ é a esperança do número de eventos contados em Y_i por unidade de N_i e representa uma taxa de ocorrência.

Um estimador consistente e não viciado de $\boldsymbol{\lambda}$ é

$$\hat{\boldsymbol{\lambda}} = \mathbf{D}_N^{-1} \mathbf{Y} = \begin{pmatrix} \frac{Y_1}{N_1} \\ \frac{Y_2}{N_2} \\ \vdots \\ \frac{Y_s}{N_s} \end{pmatrix}. \quad (4.17)$$

O vetor de esperanças e a matriz de covariâncias de $\hat{\boldsymbol{\lambda}}$ são

$$\mathbf{E}(\hat{\boldsymbol{\lambda}}) = \mathbf{E}(\mathbf{D}_N^{-1} \mathbf{Y}) = \mathbf{E} \begin{pmatrix} \frac{Y_1}{N_1} \\ \frac{Y_2}{N_2} \\ \vdots \\ \frac{Y_s}{N_s} \end{pmatrix} = \begin{pmatrix} \frac{\mu_1}{N_1} \\ \frac{\mu_2}{N_2} \\ \vdots \\ \frac{\mu_s}{N_s} \end{pmatrix} = \boldsymbol{\lambda} \quad (4.18)$$

$$\mathbf{V}_{\hat{\boldsymbol{\lambda}}}(\boldsymbol{\mu}) = \text{Var}(\mathbf{D}_N^{-1} \mathbf{Y}) = \mathbf{D}_N^{-1} [\mathbf{V}_Y(\boldsymbol{\mu})] (\mathbf{D}_N^{-1})' = \mathbf{D}_N^{-1} \mathbf{D}_\mu \mathbf{D}_N^{-1} \quad (4.19)$$

Um estimador consistente de $V_{\hat{\lambda}}(\mu)$ é obtido substituindo-se μ por \mathbf{Y} em (4.19). Esse estimador será denotado por $V_{\hat{\lambda}}$:

$$V_{\hat{\lambda}} = \mathbf{D}_N^{-1} \mathbf{D}_Y \mathbf{D}_N^{-1} = \tag{4.20}$$

$$= \begin{bmatrix} \frac{Y_1}{N_1^2} & 0 & \cdots & 0 \\ 0 & \frac{Y_2}{N_2^2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{Y_s}{N_s^2} \end{bmatrix}.$$

Segundo Plackett (1974), o estimadores não viciados Y_i de μ_i satisfazem

$$(\mu_i)^{-1/2} (Y_i - \mu_i) \xrightarrow{D} N(0, 1)$$

quando $\mu_i \rightarrow \infty$, para todo $i = 1, 2, \dots, s$. Então o vetor $\mathbf{Y} = (Y_1, Y_2, \dots, Y_s)$ satisfaz

$$(\mathbf{D}_\mu)^{-1/2} (\mathbf{Y} - \mu) \xrightarrow{D} N_s(\mathbf{0}, \mathbf{I}). \tag{4.21}$$

De (4.16) e (4.17), a expressão acima pode ser reescrita como

$$(\mathbf{D}_N)^{1/2} (\mathbf{D}_\lambda)^{-1/2} (\hat{\lambda} - \lambda) \xrightarrow{D} N_s(\mathbf{0}, \mathbf{I}). \tag{4.22}$$

quando $\mu = \mathbf{D}_N \mathbf{D}_\lambda$ cresce através de aumentos homogêneos nas unidades de exposição N_1, N_2, \dots, N_s .

De (4.22), $\hat{\lambda}$ tem distribuição aproximada normal multivariada com

$$\mathbf{E}(\hat{\lambda}) = \lambda \tag{4.23}$$

e matriz de covariâncias

$$\Sigma(\lambda) = (\mathbf{D}_N)^{-1}(\mathbf{D}_\lambda)(\mathbf{D}_N)^{-1} = \begin{bmatrix} \frac{\lambda_1}{N_1^2} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_2}{N_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_s}{N_s^2} \end{bmatrix}. \quad (4.24)$$

quando os μ_i são suficientemente grandes.

Geralmente, quando a dados são modelados pela distribuição de Poisson, o interesse é explicar a variabilidade entre componentes da esperança assintótica de funções do vetor de parâmetros de razão

$$\mathbf{F}(\hat{\lambda}) = \begin{pmatrix} F_1(\hat{\lambda}) \\ F_2(\hat{\lambda}) \\ \vdots \\ F_u(\hat{\lambda}) \end{pmatrix}$$

(λ). O modelo linear usualmente postulado é

$$\mathbf{E}_A[\mathbf{F}(\hat{\lambda})] = \mathbf{X}\beta. \quad (4.25)$$

onde \mathbf{X} é uma matriz $u \times p$ de constantes conhecidas de posto p , β é o vetor de parâmetros desconhecidos do modelo e $\mathbf{F}(\cdot)$ deve ser tal que as condições (i) e (ii) em 3.3.1 sejam satisfeitas.

A função de resposta geralmente utilizada é

$$[\mathbf{F}(\hat{\lambda})] = \ln(\hat{\lambda})$$

Segundo Koch *et al* (1984), quando todos os y_i são moderadamente grandes (≥ 10), os μ_i são considerados grandes o suficiente para permitir aproximar tais modelos por suas correspondentes expansões em séries de Taylor lineares. De um modo geral, quando $\mathbf{F}(\cdot)$ possui derivadas parciais contínuas até ordem dois,

$$\mathbf{F}(\hat{\lambda}) = \mathbf{F}(\mathbf{D}_N^{-1}\mathbf{y})$$

pode ser aproximada por sua expansão em série de Taylor linear em torno de \mathbf{y} como

$$\mathbf{F}(\mathbf{D}_N^{-1}\mathbf{y}) = \mathbf{F}(\mathbf{D}_N^{-1}\mu) + \mathbf{H}(\mu - \mathbf{y}) \quad (4.26)$$

onde \mathbf{H} é uma matriz $u \times p$ cujo elemento (i, j) é a derivada parcial de $F_i(\hat{\lambda})$ com respeito à j -ésima coordenada de $\hat{\lambda}$ ($i = 1, 2, \dots, u$ e $j = 1, 2, \dots, p$).

De (4.21) e (4.26), $\mathbf{F}(\hat{\lambda})$ tem distribuição assintótica normal multivariada com

$$\mathbf{E}_A[\mathbf{F}(\hat{\lambda})] = \mathbf{F}(\lambda) \quad (4.27)$$

e matriz de covariâncias

$$\mathbf{Var}_A[\mathbf{F}(\hat{\lambda})] = \Sigma(\mu) = \mathbf{H}(\mu)[\mathbf{V}_{\hat{\lambda}}(\mu)]\mathbf{H}'(\mu) \quad (4.28)$$

$$= \mathbf{H}(\mu)[\mathbf{D}_N^{-1}\mathbf{D}_\mu\mathbf{D}_N^{-1}]\mathbf{H}'(\mu)$$

onde $\mathbf{E}_A(\cdot)$ e $\mathbf{Var}_A(\cdot)$ denotam a esperança e a matriz de covariâncias assintóticas de $\mathbf{F}(\hat{\lambda})$, respectivamente, no sentido de expressões limite quando os y_i crescem através de aumentos homogêneos nas unidades de exposição N_i .

Substituindo μ por seu estimador consistente \mathbf{Y} em (4.28) obtêm-se um estimador consistente da matriz de covariâncias assintótica de $\mathbf{F}(\hat{\lambda})$ que será denotado por

$$\Sigma(\mathbf{Y}) = \mathbf{H}(\mathbf{Y})[\mathbf{D}_N^{-1}\mathbf{D}_Y\mathbf{D}_N^{-1}]\mathbf{H}'(\mathbf{Y}). \quad (4.29)$$

Se $\mathbf{F}(\cdot)$ é tal que $\mathbf{H}(\mathbf{Y})$ tem posto linha completo, a matriz de covariâncias obtida em (4.29) é não singular, e portanto, o método de quadrados mínimos ponderados pode ser usado na estimação de parâmetros do modelo descrito em (4.25).

Capítulo 5

Exemplos.

5.1 Exemplo 1. Associação entre peso do fruto e escurecimento interno em abacaxi Smooth Cayenne.

O problema objeto de análise nesse exemplo baseou-se num experimento descrito em Botrel & Carvalho (1983). Um dos objetivos desse experimento era avaliar a associação entre *peso do fruto e escurecimento interno* em abacaxi Smooth Cayenne, quando os frutos eram submetidos a diferentes *condições de armazenamento* (tratamentos). Os *tratamentos* consistiam de:

1. Avaliação dos frutos após 7 dias de armazenamento sem refrigeração a uma temperatura de $25^{\circ}C$ (SR);
2. Avaliação dos frutos após 15 dias de armazenamento com refrigeração em câmara climática a $5^{\circ}C$ e 7 dias em condições ambientais (CR).

A avaliação consistia em seccionar longitudinalmente os frutos e mensurar a variável resposta *percentagem da área da seção longitudinal afetada pelo escurecimento*. Foram utilizados no experimento 480 frutos, classificados de acordo com as categorias de peso especificadas na Tabela 5.1.

Nesse experimento o grau de escurecimento interno foi mensurado através de uma variável resposta contínua. A escolha da escala de mensuração da variável resposta depende do objetivo do experimento e de sua praticabilidade. Em muitas situações, a

Categoria	gramas
C1	1800 - 2300
C2	1500 - 1799
C3	1300 - 1499
C4	1100 - 1299
C5	900 - 1099
C6	700 - 899

Tabela 5.1: Categorias de peso em abacaxi Smooth Cayenne.

avaliação em escalas de mensuração inferiores permite atingir os objetivos do experimento. O grau de escurecimento poderia então ser avaliado de acordo com os seguintes escalas de mensuração:

1. Variável resposta nominal dicotômica: observa-se apenas a presença ou ausência de escurecimento interno;
2. Variável resposta ordinal politômica: os frutos são classificados visualmente em muito afetados, pouco afetados e não afetados pelo escurecimento interno.
3. Variável resposta intervalar :os frutos são classificados em intervalos com base na percentagem de área afetada pelo escurecimento.

Os dados aqui utilizados considerando essas escalas de mensuração são fictícios, e conseqüentemente, os resultados de sua análise não têm nenhuma validade do ponto de vista agrônomo. Sua utilização visa mostrar a aplicação do método GSK no estudo de um problema de interesse na área de conhecimento subjacente.

Com o objetivo de enfatizar a importância da estruturação dos dados na sua modelagem probabilística esse exemplo será analisado considerando duas possíveis estruturas para os dados. Essas estruturas correspondem a dois diferentes modos de estratificação, definidas pelos níveis (ou combinação dos níveis dos *fatores*) e são descritas a seguir:

1. Estrutura I: *Subpopulações definidas pelos níveis do tratamento*. Os 480 frutos são separados em dois grupos de 240 e os tratamentos (SR e CR) são alocados aleatori-

amente a cada grupo. Após a alocação dos tratamentos, os frutos são classificados em categorias de peso. Observe que nesse caso o número de frutos em cada categoria de peso não é fixo.

2. Estrutura II: *Subpopulações definidas pelas combinações peso × tratamento*. São escolhidos 80 frutos de cada uma das seis categorias de peso. Os tratamentos (SR e CR) são então aleatorizados dentro de cada categoria de peso.

Na estrutura I, os dados serão analisados considerando apenas a escala de mensuração

1. Na estrutura II a análise será feita supondo as escalas de mensuração 1 e 2, com o objetivo de exemplificar as modelagens pelas distribuições binomial e multinomial, respectivamente

5.1.1 Análise dos dados considerando a estrutura I.

Na estrutura I, os frutos observados em cada grupo são conceitualmente representativos de subpopulações infinitas (correspondentes a cada tratamento) no sentido de equivalência com uma amostra aleatória estratificada. As tendências de cada fruto para apresentar a r -ésima categoria de resposta (categoria de peso × grau de escurecimento interno) são consideradas mutuamente independentes e as probabilidades de cada fruto apresentar a resposta peso i e grau de escurecimento k são supostas iguais dentro de cada subpopulação.

Na avaliação dos frutos observa-se, além da variável resposta (presença ou ausência de escurecimento), a covariável peso discretizada em seis categorias (Tabela 5.1).

Estrutura I e variável resposta dicotômica.

Na Tabela 5.2 são apresentados os dados resultantes da classificação cruzada dos frutos de acordo com o grau de escurecimento (com e sem escurecimento) e categoria de peso (C1, C2, ..., C6) em cada tratamento (SR e CR).

Cada linha da Tabela 5.2 corresponde a observações do vetor aleatório

Subpopulação	C1		C2		C3		C4		C5		C6	
	S	N	S	N	S	N	S	N	S	N	S	N
SR	10	30	15	25	8	32	6	34	4	36	5	35
CR	18	22	19	21	11	29	9	31	9	31	7	33

Tabela 5.2: Número de frutos com e sem escurecimento interno em cada tratamento (SR e CR) e categoria de peso (C1, C2, ..., C6).

$$\mathbf{Y}_i = (Y_{i11}, Y_{i12}, Y_{i21}, Y_{i22}, \dots, Y_{ijk}, \dots, Y_{i61}, Y_{i62})$$

onde Y_{ijk} é o número de frutos na j -ésima categoria de peso ($j = 1, 2, \dots, 6$) com a resposta k no i -ésimo tratamento ($i = 1, 2$). Nesse caso, o número de *subpopulações* é igual ao número de tratamentos ($s = 2$) e as *categorias de resposta* resultam do cruzamento dos $j = 6$ níveis da covariável peso com os $k = 2$ níveis variável resposta ($r = 12$).

Considera-se que \mathbf{Y}_i tem distribuição multinomial com função de probabilidade

$$P(Y_{i11} = y_{i11}, \dots, Y_{i62} = y_{i62}) = \frac{n_i}{\prod_{j=1}^6 \prod_{k=1}^2 y_{ijk}!} \prod_{j=1}^6 \prod_{k=1}^2 \pi_{ijk}^{y_{ijk}}$$

onde $n_i = 240$ é o tamanho da amostra em cada subpopulação, y_{ijk} é o valor observado de Y_{ijk} e π_{ijk} denota a probabilidade de que um fruto selecionado aleatoriamente na subpopulação correspondente ao tratamento i seja classificada na categoria de resposta jk . Observe que $\sum_{j=1}^6 \sum_{k=1}^2 y_{ijk} = n_i$ e $\sum_{j=1}^6 \sum_{k=1}^2 \pi_{ijk} = 1$ para todo i .

Dessa forma o vetor aleatório combinado

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

é modelado pela distribuição *produto de multinomiais* com função de probabilidade

$$P(Y_{111} = y_{111}, \dots, Y_{362} = y_{362}) = \prod_{i=1}^3 \frac{n_i}{\prod_{j=1}^6 \prod_{k=1}^2 y_{ijk}!} \prod_{j=1}^6 \prod_{k=1}^2 \pi_{ijk}^{y_{ijk}} \quad (5.1)$$

com y_{ijk} , n_i e π_{ijk} conforme definidos anteriormente.

O vetor de proporções amostrais

$$\mathbf{p} = (p_{111}, p_{112}, \dots, p_{161}, p_{211}, p_{212}, \dots, p_{261})$$

é um estimador consistente e não viciado do vetor de parâmetros π que indexa a família de distribuições de \mathbf{Y} .

Esperança e covariância de \mathbf{p} . De (4.4) e (4.5), o vetor de esperanças de \mathbf{p} é

$$\mathbf{E}(\mathbf{p}) = \begin{pmatrix} E(p_{111}) \\ E(p_{112}) \\ \vdots \\ E(p_{162}) \\ E(p_{211}) \\ E(p_{212}) \\ \vdots \\ E(p_{262}) \end{pmatrix} = \begin{pmatrix} \pi_{111} \\ \pi_{112} \\ \vdots \\ \pi_{162} \\ \pi_{211} \\ \pi_{212} \\ \vdots \\ \pi_{262} \end{pmatrix} \quad (5.2)$$

e sua matriz de covariâncias, uma matriz bloco diagonal 24×24

$$\mathbf{V}_{\mathbf{p}}(\pi) = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} \quad (5.3)$$

onde

$$V_i = \frac{1}{n_i} \begin{bmatrix} \pi_{i11}(1 - \pi_{i11}) & -\pi_{i11}\pi_{i12} & \cdot & \cdot & \cdot & -\pi_{i11}\pi_{i62} \\ -\pi_{i11}\pi_{i12} & \pi_{i12}(1 - \pi_{i12}) & \cdot & \cdot & \cdot & -\pi_{i12}\pi_{i62} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -\pi_{i11}\pi_{i61} & -\pi_{i12}\pi_{i62} & \cdot & \cdot & \cdot & \pi_{i62}(1 - \pi_{i62}) \end{bmatrix}$$

Observa-se que cada bloco é a matriz de covariâncias do subvetor de proporções amostrais correspondente à população i

$$p_i = (p_{i11}, p_{i12}, \dots, p_{i62}).$$

Como as subpopulações são independentes, a covariância entre dois elementos quaisquer de p pertencentes a subpopulações diferentes é zero.

Substituindo-se π por p em (5.2) e (5.3) obtêm-se estimadores da esperança e matriz de covariâncias assintótica de p .

Escolha da função de resposta. Para investigar a associação entre peso e escurecimento interno podem ser formuladas as seguintes hipóteses científicas:

1. Existe associação entre peso e escurecimento interno quando os frutos são armazenados sem refrigeração (tratamento 1);
2. Existe associação entre peso e escurecimento interno quando os frutos são armazenados com refrigeração (tratamento 2);
3. A natureza da associação é igual nos dois tratamentos.

As hipóteses científicas acima definidas serão representadas por hipóteses estatísticas, isto é, restrições que o vetor de parâmetros π deve satisfazer. Se o interesse do pesquisador é investigar a existência de uma associação multiplicativa, os conjuntos de restrições sobre π para as hipóteses científicas 1 e 2 podem ser representados como

$$H1_0 : \frac{\pi_{111}}{\pi_{112}} = \frac{\pi_{121}}{\pi_{122}} = \dots = \frac{\pi_{161}}{\pi_{162}} \quad (5.4)$$

e

$$H2_0 : \frac{\pi_{211}}{\pi_{212}} = \frac{\pi_{221}}{\pi_{222}} = \dots = \frac{\pi_{261}}{\pi_{262}}, \quad (5.5)$$

respectivamente. A ausência de associação multiplicativa entre peso e escurecimento interno significa, portanto, que a razão entre a proporção de frutos com e sem escurecimento em cada categoria de peso é constante. Observe que numa hipótese de não *associação multiplicativa* são comparadas razões entre probabilidades, enquanto numa hipótese de não *associação aditiva* são consideradas diferenças (Forthofer & Lehnen (1981)).

A hipótese de homogeneidade da associação definida em 3 tem a seguinte representação:

$$H3_0 : \frac{\pi_{111}}{\pi_{112}} \times \frac{\pi_{212}}{\pi_{211}} = \frac{\pi_{121}}{\pi_{122}} \times \frac{\pi_{222}}{\pi_{221}} = \dots = \frac{\pi_{161}}{\pi_{162}} \times \frac{\pi_{262}}{\pi_{261}}. \quad (5.6)$$

$H1_0, H2_0$ e $H3_0$, do modo que foram representadas acima, não constituem hipóteses lineares sobre π e dessa forma não podem ser testadas pela estatística de Wald.

Então, a função $\mathbf{F}(\cdot)$ deve ser definida de modo que as hipóteses de interesse do estudo ($H1_0, H2_0$ e $H3_0$) possam ser reescritas com restrições lineares sobre $\mathbf{F}(\pi)$. Além disso, $\mathbf{F}(\cdot)$, deve satisfazer às condições (i) e (ii) descritas na seção 3.3.1.

Definimos então $\mathbf{F}(\mathbf{p})$ como

$$\mathbf{F} : IR^{24} \mapsto IR^{12}$$

com

$$\begin{aligned}
 \mathbf{F}(\mathbf{p}) &= \begin{pmatrix} F_{11} \\ F_{12} \\ \vdots \\ F_{16} \\ F_{21} \\ F_{22} \\ \vdots \\ F_{26} \end{pmatrix} = \begin{pmatrix} \ln\left(\frac{p_{111}}{p_{112}}\right) \\ \ln\left(\frac{p_{121}}{p_{122}}\right) \\ \vdots \\ \ln\left(\frac{p_{161}}{p_{162}}\right) \\ \ln\left(\frac{p_{211}}{p_{212}}\right) \\ \ln\left(\frac{p_{221}}{p_{222}}\right) \\ \vdots \\ \ln\left(\frac{p_{261}}{p_{262}}\right) \end{pmatrix} = \quad (5.7) \\
 &= \begin{pmatrix} \ln(p_{111}) - \ln(p_{112}) \\ \ln(p_{121}) - \ln(p_{122}) \\ \vdots \\ \ln(p_{161}) - \ln(p_{162}) \\ \ln(p_{211}) - \ln(p_{212}) \\ \ln(p_{221}) - \ln(p_{222}) \\ \vdots \\ \ln(p_{261}) - \ln(p_{262}) \end{pmatrix}
 \end{aligned}$$

onde cada F_{ij} é o logaritmo natural da razão entre a proporção de frutos com e sem escurecimento no tratamento i e categoria de peso j .

Distribuição assintótica de $\mathbf{F}(\mathbf{p})$. Para calcular a distribuição assintótica de $\mathbf{F}(\mathbf{p})$ e conseqüentemente sua esperança e matriz de covariâncias assintóticas será usado o método delta multivariado.

De (4.10), o vetor de proporções amostrais \mathbf{p} é uma estatística que satisfaz

$$\sqrt{n}(\mathbf{p} - \boldsymbol{\pi}) \xrightarrow{\mathcal{D}} N_{24}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{p}}(\boldsymbol{\pi})) \quad (5.8)$$

onde n é o número de frutos utilizados em cada tratamento ($n_1 = n_2 = n$) e $\boldsymbol{\Sigma}_{\mathbf{p}}(\boldsymbol{\pi})$ é a matriz de covariâncias assintótica de \mathbf{p} ,

$$\Sigma_{\mathbf{p}}(\pi) = n \mathbf{V}_{\mathbf{p}}(\pi)$$

com $\mathbf{V}_{\mathbf{p}}(\pi)$ definida em (5.3).

Sejam $\mathbf{H} = \frac{\partial \mathbf{F}}{\partial \mathbf{p}}$ e $\mathbf{G} = \frac{\partial^2 \mathbf{F}}{\partial^2 \mathbf{p}}$ matrizes cujos elementos (k, l) são as derivadas de ordem 1 e 2, respectivamente, da k -ésima componente de $\mathbf{F} = (F_{11}(\mathbf{p}), F_{12}(\mathbf{p}), \dots, F_{26}(\mathbf{p})) \equiv (F_1, F_2, \dots, F_k, \dots, F_{12})$, com respeito ao l -ésimo elemento de $\mathbf{p} = (p_{111}, p_{112}, \dots, p_{162}, p_{211}, p_{212}, \dots, p_{262}) \equiv (p_1, p_2, \dots, p_{24})$. Assim, $\frac{\partial \mathbf{F}}{\partial \mathbf{p}}$ é uma matriz bloco diagonal 12×24

$$\mathbf{H} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

onde D_i , para $i = 1, 2$ é da forma:

$$D_i = \begin{bmatrix} \frac{\partial F_{i1}}{\partial p_{i11}} & \frac{\partial F_{i1}}{\partial p_{i12}} & \dots & \frac{\partial F_{i2}}{\partial p_{i62}} \\ \frac{\partial F_{i2}}{\partial p_{i11}} & \frac{\partial F_{i2}}{\partial p_{i12}} & \dots & \frac{\partial F_{i2}}{\partial p_{i62}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_{i6}}{\partial p_{i11}} & \frac{\partial F_{i6}}{\partial p_{i12}} & \dots & \frac{\partial F_{i6}}{\partial p_{i62}} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{p_{i11}} & \frac{-1}{p_{i12}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{p_{i21}} & \frac{-1}{p_{i22}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{p_{i31}} & \frac{-1}{p_{i3211}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{p_{i41}} & \frac{-1}{p_{i42}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{p_{i51}} & \frac{-1}{p_{i52}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{p_{i61}} & \frac{-1}{p_{i62}} \end{bmatrix} \quad (5.9)$$

De modo semelhante, a matriz \mathbf{G} de derivadas parciais de segunda ordem, é expressa por

$$\mathbf{G} = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}$$

onde C_i , para $i = 1, 2$ é:

$$C_i = \begin{bmatrix} \frac{\partial F_{i1}^2}{\partial p_{i11}^2} & \frac{\partial F_{i1}^2}{\partial p_{i12}^2} & \dots & \frac{\partial F_{i1}^2}{\partial p_{i\epsilon 2}^2} \\ \frac{\partial F_{i2}^2}{\partial p_{i11}^2} & \frac{\partial F_{i2}^2}{\partial p_{i12}^2} & \dots & \frac{\partial F_{i2}^2}{\partial p_{i\epsilon 2}^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_{i\epsilon}^2}{\partial p_{i11}^2} & \frac{\partial F_{i\epsilon}^2}{\partial p_{i12}^2} & \dots & \frac{\partial F_{i\epsilon}^2}{\partial p_{i\epsilon 2}^2} \end{bmatrix} = \quad (5.10)$$

$$= \begin{bmatrix} \frac{-1}{p_{i11}^2} & \frac{1}{p_{i12}^2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-1}{p_{i21}^2} & \frac{1}{p_{i22}^2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-1}{p_{i31}^2} & \frac{1}{p_{i32}^2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{-1}{p_{i41}^2} & \frac{-1}{p_{i42}^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{-1}{p_{i51}^2} & \frac{1}{p_{i52}^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{-1}{p_{i\epsilon 1}^2} & \frac{1}{p_{i\epsilon 2}^2} \end{bmatrix}.$$

De (5.9) e (5.10) observa-se que $\mathbf{F}(\mathbf{p})$ possui derivadas contínuas até ordem 2 para todo $\mathbf{p} \in (0, 1)^{24}$, e conseqüentemente numa região contendo π . Satisfeita essa propriedade, $\mathbf{F}(\mathbf{p})$ pode ser expandida segundo (3.6).

Por (4.7) e (4.6) e aplicando o teorema que resume o método delta multivariado, a distribuição assintótica de $\mathbf{F}(\mathbf{p})$ é dada por

$$\sqrt{n}[\mathbf{F}(\mathbf{p}) - \mathbf{F}(\pi)] \xrightarrow{D} N_{24}[0, \mathbf{H}(\pi)\Sigma_{\mathbf{p}}(\pi)\mathbf{H}(\pi)']. \quad (5.11)$$

ou seja, para n grande, $\mathbf{F}(\mathbf{p})$ tem distribuição aproximada normal multivariada com esperança $\mathbf{F}(\pi)$ e matriz de covariâncias $n^{-1}\mathbf{H}(\pi)\Sigma_{\mathbf{p}}(\pi)\mathbf{H}(\pi)'$. Para obter um estimador consistente da matriz de covariâncias assintótica de $\mathbf{F}(\mathbf{p})$,

$$\Sigma_{\mathbf{F}(\mathbf{p})}(\pi) = \mathbf{H}(\pi)\Sigma_{\mathbf{p}}(\pi)\mathbf{H}(\pi)' \quad (5.12)$$

substituímos π por p em (5.12).

Definição do modelo.

A parametrização que define o modelo linear geral a ser postulado foi escolhida de modo que as hipóteses de interesse possam ser escritas como restrições lineares sobre esses parâmetros e além disso, a matriz do modelo correspondente a essa parametrização tenha posto completo.

Observe que cada coordenada corresponde ao logarítmo natural da razão entre a proporção de frutos com e sem escurecimento em cada tratamento e categoria de peso. Se postularmos o modelo identidade

$$E_A[F(p)] = X\beta \quad (5.13)$$

onde X é uma matriz identidade 12×12 , as componentes do vetor de parâmetros $\beta = (\beta_{11}, \beta_{12}, \dots, \beta_{25}, \beta_{26})$ coincidem com as componentes de $F(\pi)$, sob o modelo. Dessa forma, os parâmetros β_{ij} ($i = 1, 2$ e $j = 1, \dots, 6$) correspondem ao logarítmo natural da proporção de frutos com e sem escurecimento na categoria de peso j , quando submetidos ao tratamento ($i = 1, 2$ e $j = 1, \dots, 6$) e as hipóteses definidas em (5.4), (5.5) e (5.6) podem ser postuladas como restrições lineares sobre β .

O conjunto de restrições não lineares sobre π representado em H_{10} pode ser reescrito como

$$H_{10} : \ln\left(\frac{\pi_{111}}{\pi_{112}}\right) = \ln\left(\frac{\pi_{121}}{\pi_{122}}\right) = \dots = \ln\left(\frac{\pi_{161}}{\pi_{162}}\right)$$

que corresponde aos seguintes contrastes linearmente independentes:

$$\beta_{11} - \beta_{12} = 0$$

$$\beta_{11} - \beta_{13} = 0$$

$$\beta_{11} - \beta_{14} = 0$$

$$\beta_{11} - \beta_{15} = 0$$

$$\beta_{11} - \beta_{16} = 0$$

Sua representação matricial é

$$H1_0 : C_1 \beta = 0$$

onde C_1 é uma matriz 5×12 de posto-linha completo da seguinte forma

$$C_1 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

De modo semelhante $H2_0$ pode ser reescrita como

$$H1_0 : C_2 \beta = 0$$

onde

$$C_2 = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

e $H3_0$ como

$$H3_0 : C_3\beta = 0$$

onde

$$C_3 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Análise dos resultados.

Os valores observados da função de resposta e seus respectivos erros padrão são apresentados na Tabela 5.3. No modelo definido em (5.13) (modelo identidade saturado) as estimativas dos parâmetros correspondem aos valores observados da função, pois

$$E_A(F) = I\beta = \beta.$$

Dessa forma, o valor estimado de cada parâmetro, corresponde ao logaritmo natural da razão entre as proporções de frutos com e sem escurecimento para cada tratamento e categoria de peso. Os valores negativos em todas as estimativas (Tabela 5.3) indicam que a proporção de frutos com escurecimento foi sempre inferior à proporção sem escurecimento.

A análise dos contrastes correspondentes às hipóteses $H1_0$, $H2_0$ e $H3_0$ (Tabela 5.4) indica a existência de associação entre *peso do fruto* e *escurecimento interno* nos tratamentos *sem refrigeração* ($p = 0,0446$) e *com refrigeração* ($p = 0,0076$). Não houve evidência suficiente para rejeitar a hipótese de homogeneidade da associação ($p = 0,8932$), isto é, a natureza da associação multiplicativa entre peso e escurecimento interno é a mesma quando os frutos são armazenados com e sem refrigeração.

FUNÇÃO	ESTIMATIVA	ERRO PADRÃO
SRC1	-1,2368	0,3786
SRC2	-0,4700	0,3291
SRC3	-1,4171	0,3941
SRC4	-1,6740	0,4449
SRC5	-2,2246	0,5263
SRC6	-1,7636	0,4419
CRC1	-0,2007	0,3178
CRC2	-0,0488	0,3124
CRC3	-1,0647	0,3667
CRC4	-1,1631	0,3623
CRC5	-1,3545	0,3965
CRC6	-1,5198	0,4172

Tabela 5.3: Valores observados da função de resposta em cada tratamento e categoria de peso com seus respectivos erros padrão.

CONTRASTE	G.L.	QUI-QUADRADO	PROB
PESO X ESC EM SR	5	11,37	0,0446
PESO X ESC EM CR	5	15,76	0,0076
HOMOGENEIDADE	5	1,67	0,8932

Tabela 5.4: Contrastes correspondentes às hipóteses H_{10} , H_{20} e H_{30} com respectivos graus de liberdade, valores observados das estatísticas de teste e valores de p associados.

5.1.2 Análise dos dados considerando a estrutura II.

Na estrutura II, a estratificação por peso é feita *a priori* e as combinações das categorias de peso ($C1, \dots, C6$) com os tratamentos (SR e CR) correspondem a s subpopulações ($s = 12$). Nesse caso, as categorias de resposta coincidem com o número de níveis da variável resposta. Os dados serão analisados supondo resposta dicotômica (*com escurecimento e sem escurecimento*) e resposta ordinal politômica (níveis de escurecimento *alto, baixo e nulo*).

Os frutos avaliados nesse experimento são considerados conceitualmente representativos de subpopulações infinitas no sentido de equivalência com uma amostra aleatória estratificada onde os estratos correspondem cada grupo (categoria de peso \times tratamento).

As tendências de cada fruto para apresentar a resposta k são consideradas mutuamente independentes. Além disso a probabilidade de um fruto selecionado ao acaso apresentar a resposta k é contante dentro da mesma subpopulação.

Estrutura II e variável resposta dicotômica.

Y_{ij} é modelado pela distribuição binomial com parâmetros $n = 120$ e π_{ij} . A função de probabilidade de Y_{ij} é

$$P(Y_{ij} = y_{ij}) = \binom{n}{y_{ij}} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{n - y_{ij}}$$

onde $n = 40$ é o tamanho da amostra em cada subpopulação. y_{ij} é o valor observado de Y_{ij} e π_{ij} denota a probabilidade de que um fruto selecionado aleatoriamente na subpopulação correspondente à categoria de peso i e tratamento j apresente escurecimento interno.

Os resultados observados para cada subpopulação são apresentados numa tabela de contingência 12×2 (Tabela 5.5). Esses dados correspondem a observações do vetor aleatório

$$\mathbf{Y} = (Y_{11}, Y_{12}, Y_{21}, Y_{22}, \dots, Y_{61}, Y_{62}).$$

$Y_{11}, Y_{12}, Y_{21}, Y_{22}, \dots, Y_{61}, Y_{62}$ são variáveis aleatórias independentes com $Y_{ij} \sim B(n, \pi_{ij})$,

SUBPOPULAÇÃO	S	N
C1 SR	10	30
C1 CR	18	22
C2 SR	15	25
C2 CR	19	21
C3 SR	8	32
C3 CR	11	29
C4 SR	6	34
C4 CR	9	31
C5 SR	4	36
C5 CR	9	31
C6 SR	5	35
C6 CR	7	33

Tabela 5.5: Número de frutos com (S) e sem (N) escurecimento interno em cada categoria de peso (C1, C2, ... , C6) e tratamento (SR, CR).

logo \mathbf{Y} tem distribuição produto de binomiais com função de probabilidade

$$P(Y_{11} = y_{11}, \dots, Y_{62} = y_{62}) = \prod_{i=1}^6 \prod_{j=1}^2 \binom{n}{y_{ij}} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{n-y_{ij}}$$

com n , y_{ij} , e π_{ij} conforme definidos anteriormente.

Um estimador consistente e não viciado de $\boldsymbol{\pi} = (\pi_{11}, \pi_{12}, \dots, \pi_{62})$ é o vetor de proporções amostrais $\mathbf{p} = (p_{11}, p_{12}, \dots, p_{62})$.

Esperança e covariância de \mathbf{p} .

De (4.4) e (4.5), o vetor de esperanças e a matriz de covariâncias de \mathbf{p} são

$$\mathbf{E}(\mathbf{p}) = \begin{pmatrix} E(p_{11}) \\ E(p_{12}) \\ \vdots \\ E(p_{61}) \\ E(p_{62}) \end{pmatrix} = \begin{pmatrix} \pi_{11} \\ \pi_{12} \\ \vdots \\ \pi_{61} \\ \pi_{62} \end{pmatrix}$$

e

$$\mathbf{V}_p(\pi) = \frac{1}{n} \begin{bmatrix} \pi_{11}(1 - \pi_{11}) & 0 & \dots & 0 \\ 0 & \pi_{12}(1 - \pi_{12}) & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & \pi_{62}(1 - \pi_{62}) \end{bmatrix} \quad (5.14)$$

respectivamente.

As componentes de \mathbf{p} são linearmente independentes. Assim, $\mathbf{V}_p(\pi)$ é uma matriz 12×12 positiva definida e portanto de posto completo. Um estimador consistente de $\mathbf{V}_p(\pi)$, denotado por \mathbf{V}_p , é obtido substituindo-se π por \mathbf{p} em (5.14):

$$\mathbf{V}_p = \frac{1}{n} \begin{bmatrix} p_{11}(1 - p_{11}) & 0 & \dots & 0 \\ 0 & p_{12}(1 - p_{12}) & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & p_{61}(1 - p_{62}) \end{bmatrix}.$$

Escolha da função de resposta. Suponha que seja de interesse testar a hipótese de ausência de *interação multiplicativa* entre *peso* e *tratamento*. Tal hipótese pode ser representada por

$$H1_0 : \frac{\pi_{11}}{1 - \pi_{11}} \times \frac{1 - \pi_{12}}{\pi_{12}} = \frac{\pi_{21}}{1 - \pi_{21}} \times \frac{1 - \pi_{22}}{\pi_{22}} = \dots = \frac{\pi_{61}}{1 - \pi_{61}} \times \frac{1 - \pi_{62}}{\pi_{62}} \quad (5.15)$$

Define-se então $\mathbf{F}(\mathbf{p})$ de modo que $H1_0$ possa ser representada por um conjunto de

restrições lineares sobre π . Seja

$$\mathbf{F} : \mathbb{R}^{24} \longrightarrow \mathbb{R}^{12}$$

com

$$\begin{aligned} \mathbf{F}(\mathbf{p}) &= \begin{pmatrix} F_{11} \\ F_{12} \\ F_{21} \\ F_{22} \\ \vdots \\ F_{61} \\ F_{62} \end{pmatrix} = \begin{pmatrix} \ln\left(\frac{p_{11}}{1-p_{11}}\right) \\ \ln\left(\frac{p_{12}}{1-p_{12}}\right) \\ \ln\left(\frac{p_{21}}{1-p_{21}}\right) \\ \ln\left(\frac{p_{22}}{1-p_{22}}\right) \\ \vdots \\ \ln\left(\frac{p_{61}}{1-p_{61}}\right) \\ \ln\left(\frac{p_{62}}{1-p_{62}}\right) \end{pmatrix} = \\ &= \begin{pmatrix} \ln(p_{11}) - \ln(1-p_{11}) \\ \ln(p_{12}) - \ln(1-p_{12}) \\ \ln(p_{21}) - \ln(1-p_{21}) \\ \ln(p_{22}) - \ln(1-p_{22}) \\ \vdots \\ \ln(p_{61}) - \ln(1-p_{61}) \\ \ln(p_{62}) - \ln(1-p_{62}) \end{pmatrix} \end{aligned} \quad (5.16)$$

onde cada F_{ij} é o logaritmo natural da razão entre a proporção de frutos com e sem escurecimento na categoria de peso i e tratamento j .

A hipótese definida em $H1_0$ pode então ser reescrita como

$$H1_0 : F_{11}(\pi) - F_{12}(\pi) = F_{21}(\pi) - F_{22}(\pi) = \dots = F_{61}(\pi) - F_{62}(\pi).$$

Sua representação matricial é

$$H1_0 : \mathbf{C}_1 \mathbf{F}(\pi) = \mathbf{0} \quad (5.17)$$

onde

$$C_1 = \begin{bmatrix} 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

Distribuição assintótica de $F(\mathbf{p})$. Para calcular a distribuição assintótica de $F(\mathbf{p})$ e conseqüentemente sua esperança e matriz de covariâncias assintóticas será usado o método delta multivariado.

De (4.10),

$$\sqrt{n}(\mathbf{p} - \pi) \xrightarrow{D} N_{24}(\mathbf{0}, \Sigma_{\mathbf{p}}(\pi)) \quad (5.18)$$

onde n é tamanho da amostra em cada subpopulação ($n_{11} = n_{12} = \dots = n_{62} = n$) e

$$\Sigma_{\mathbf{p}}(\pi) = n \mathbf{V}_{\mathbf{p}}(\pi) \quad (5.19)$$

com $\mathbf{V}_{\mathbf{p}}(\pi)$ definida em (5.14).

Sejam $\mathbf{H} = \frac{\partial \mathbf{F}}{\partial \mathbf{p}}$ e $\mathbf{G} = \frac{\partial^2 \mathbf{F}}{\partial^2 \mathbf{p}}$ matrizes cujos elementos (k, l) são as derivadas de ordem 1 e 2, respectivamente, da k -ésima componente de $\mathbf{F} = (F_{11}(\mathbf{p}), F_{12}(\mathbf{p}), \dots, F_{62}(\mathbf{p})) \equiv (F_1, F_2, \dots, F_k, \dots, F_{12})$, com respeito ao l -ésimo elemento de $\mathbf{p} = (p_{11}, p_{12}, \dots, p_{62}) = (p_1, p_2, \dots, p_{12})$.

Assim, \mathbf{H} é uma matriz diagonal 12×12 da forma:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial F_{11}}{\partial p_{11}} & \frac{\partial F_{11}}{\partial p_{12}} & \dots & \frac{\partial F_{11}}{\partial p_{62}} \\ \frac{\partial F_{12}}{\partial p_{11}} & \frac{\partial F_{12}}{\partial p_{12}} & \dots & \frac{\partial F_{12}}{\partial p_{62}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_{62}}{\partial p_{11}} & \frac{\partial F_{62}}{\partial p_{12}} & \dots & \frac{\partial F_{62}}{\partial p_{62}} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{p_{11}(1-p_{11})} & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{p_{11}(1-p_{11})} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{p_{11}(1-p_{11})} & 0 \\ 0 & 0 & \cdots & 0 & \frac{1}{p_{11}(1-p_{11})} \end{bmatrix} \quad (5.20)$$

De modo semelhante, a matriz \mathbf{G} de derivadas parciais de segunda ordem é expressa por

$$\mathbf{G} = \begin{bmatrix} \frac{\partial \mathbf{F}_{11}^2}{\partial p_{11}^2} & \frac{\partial \mathbf{F}_{11}^2}{\partial p_{12}^2} & \cdots & \frac{\partial \mathbf{F}_{11}^2}{\partial p_{62}^2} \\ \frac{\partial \mathbf{F}_{12}^2}{\partial p_{11}^2} & \frac{\partial \mathbf{F}_{12}^2}{\partial p_{12}^2} & \cdots & \frac{\partial \mathbf{F}_{12}^2}{\partial p_{62}^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{F}_{62}^2}{\partial p_{11}^2} & \frac{\partial \mathbf{F}_{62}^2}{\partial p_{12}^2} & \cdots & \frac{\partial \mathbf{F}_{62}^2}{\partial p_{62}^2} \end{bmatrix} = \begin{bmatrix} \frac{-1}{p_{11}^2} + \frac{1}{(1-p_{11})^2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{-1}{p_{12}^2} + \frac{1}{(1-p_{12})^2} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{-1}{p_{61}^2} + \frac{1}{(1-p_{61})^2} & 0 \\ 0 & 0 & \cdots & 0 & \frac{-1}{p_{62}^2} + \frac{1}{(1-p_{62})^2} \end{bmatrix}. \quad (5.21)$$

De (5.20) e (5.21) observa-se que $\mathbf{F}(\mathbf{p})$ possui derivadas contínuas até ordem 2 para todo $\mathbf{p} \in (0,1)^{12}$, e conseqüentemente numa região contendo π . Satisfeita essa propriedade, por (4.6), $\mathbf{F}(\mathbf{p})$ tem a seguinte expansão quando $\mathbf{p} \rightarrow \pi$:

$$\mathbf{F}(\mathbf{p}) = \mathbf{F}(\pi) + \mathbf{H}(\mathbf{p} - \pi) + o(\|\mathbf{p} - \pi\|) \quad (5.22)$$

Por (5.18) e (5.22), (4.7) vale. Então a distribuição assintótica de $\mathbf{F}(\mathbf{p})$ é dada por

$$\sqrt{n}[\mathbf{F}(\mathbf{p}) - \mathbf{F}(\pi)] \xrightarrow{\mathcal{D}} N_{12}[0, \mathbf{H}(\pi)\Sigma_{\mathbf{p}}(\pi)\mathbf{H}(\pi)'] \quad (5.23)$$

ou seja, para n grande, $\mathbf{F}(\mathbf{p})$ tem distribuição aproximada normal multivariada com

esperança $F(\pi)$ e matriz de covariâncias $n^{-1}H(\pi)\Sigma_p(\pi)H(\pi)'$ com $\Sigma_p(\pi)$ definida em (5.19). Para obter um estimador consistente da matriz de covariâncias assintótica de $F(p)$,

$$\Sigma_{F(p)}(\pi) = H(\pi)\Sigma_p(\pi)H(\pi)' \quad (5.24)$$

substituímos π por p em (5.24).

Definição do modelo. Para explicar a variabilidade entre as componentes de $F(\pi)$ foi proposto o modelo de efeitos principais com interação

$$E_A(F) = X_1\beta_1 \quad (5.25)$$

onde

$$X_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 \\ 1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

e as componentes do vetor de parâmetros têm a seguinte interpretação:

β_0 :média geral;

β_1 a β_5 :efeitos diferenciais das categorias de peso C1 a C5, respectivamente;

β_6 :efeito diferencial do tratamento SR;

SUBPOPULAÇÃO	FUNÇÃO	ERRO PADRÃO
C1 SR	-1,0986	0.3651
C1 CR	-0,2008	0.3178
C2 SR	-1,3863	0.3953
C2 CR	-0,9694	0.3541
C3 SR	-2,1972	0.5270
C3 CR	-1,2368	0.3786
C4 SR	-0,5108	0.3266
C4 CR	-0,1001	0.3166
C5 SR	-1,7346	0.4428
C5 CR	-1,2368	0.3786
C6 SR	-1,9459	0.4781
C6 CR	-1,5506	0.4161

Tabela 5.6: Logitos observados e seus respectivos erros padrão para cada categoria de peso (C1, C2, ... , C6) e tratamento (SR, CR).

β_7 a β_{11} :efeitos das interações das categorias de peso C1 a C5 com o tratamento SR, respectivamente.

Análise dos resultados. Os valores dos logitos observados e seus respectivos erros-padrão são apresentados na Tabela 5.6. O qui-quadrado correspondente à fonte de variação interação peso \times tratamento na tabela de análise de variância (Tabela 5.7) do modelo (5.25) ($Q_C = 1, 11$) equivale à estatística do teste da hipótese de ausência de interação definida em (5.15). Não havendo evidência para rejeição dessa hipótese ($p = 0, 9531$), foi proposto o modelo de efeitos principais

$$E_A(F) = X_2\beta_2 \quad (5.26)$$

onde X_2 corresponde às sete primeiras colunas do modelo anterior e β_2 equivale a β_1 excluindo-se as componentes relativas às interações.

O ajuste do modelo (5.25) é dado pelo qui-quadrado do resíduo ($Q_W = 1, 11$) na tabela de análise de variância do modelo de efeitos principais (Tabela 5.8). Observe que esse valor é igual ao valor observado da estatística do teste da hipótese de não interação. O

FONTE	G.L.	QUI-QUADRADO	PROB
INTERCEPTO	1	106,47	0,0000
PESO	5	24,99	0.0001
TRATAMENTO	1	6,80	0.0091
PESO x TRATAMENTO	5	1,11	0.9531
RESIDUO	0	.	.

Tabela 5.7: Análise de variância do modelo de efeitos principais com interação.

FONTE	G.L.	QUI-QUADRADO	PROB
INTERCEPTO	1	106,47	0,0000
PESO	5	24,99	0.0001
TRATAMENTO	1	6,80	0.0091
RESÍDUO	5	1,11	0.9531

Tabela 5.8: Análise de variância do modelo de efeitos principais.

efeito diferencial do tratamento *sem refrigeração* (SR) foi estatisticamente significativo ($p = 0,0091$). Na Tabela 5.9 são apresentadas as estimativas de quadrados mínimos ponderados dos parâmetros do modelo (5.26) e seus respectivos erros-padrão.

Observa-se que os efeitos diferenciais das categorias de peso 2 e 5 (β_2 e β_5) são estatisticamente nulos. Dessa forma, pode ser ajustado um novo modelo retirando-se os parâmetros β_2 e β_5 e as correspondentes colunas na matriz X_2 :

$$E_A(F) = X_3\beta_3 \quad (5.27)$$

A interpretação das componentes do vetor de parâmetros $\beta_3 = (\beta_0, \beta_1, \beta_3, \beta_4, \beta_6)$ é a mesma descrita no modelo de efeitos principais com interação.

A análise de variância do modelo (5.27), as estimativas dos parâmetros e os valores preditos da função de resposta são apresentados nas tabelas 5.10, 5.11 e 5.12, respectivamente.

O valor de $p = 0,9096$ (Tabela 5.10) correspondente à estatística Q_W do modelo (5.27)

PARÂMETRO	ESTIMATIVA	ERRO PADRÃO	QUI-QUADRADO	PROB
β_0	-1,1726	0,1136	106,60	0,0000
β_1	0,5443	0,2257	5,82	0,0159
β_2	-0,0146	0,2429	0,00	0,9519
β_3	-0,4850	0,2757	3,09	0,0786
β_4	0,8644	0,2174	15,81	0,0001
β_5	-0,3201	0,2605	1,51	0,2191
β_6	-0,2939	0,1105	7,07	0,0078

Tabela 5.9: Estimativas de quadrados mínimos ponderados dos parâmetros do modelo de efeitos principais e seus respectivos erros padrão, estatísticas dos testes qui-quadrado e valores de p associados.

FONTE	G.L.	QUI-QUADRADO	PROB
MODELO	4	29.32	0.0000
RESÍDUO	7	2,72	0.9096

Tabela 5.10: Análise de variância do modelo reduzido.

PARÂMETRO	ESTIMATIVA	ERRO PADRÃO	QUI-QUADRADO	PROB
β_0	-1,1681	0,1134	106,08	0,0000
β_2	0,4796	0,2174	4,87	0,0274
β_3	-0,5886	0,2571	5,24	0,0220
β_4	0,8057	0,2104	14,66	0,0001
β_6	-0,2942	0,1105	7,08	0,0078

Tabela 5.11: Estimativas de quadrados mínimos ponderados dos parâmetros do modelo reduzido e seus respectivos erros padrão, estatísticas dos testes qui-quadrado e valores de p associados.

SUBPOPULAÇÃO	FUNÇÃO	ERRO PADRÃO
C1 SR	-0,9827	0.2635
C1 CR	-0,3944	0.2511
C2 SR	-1,4623	0.1694
C2 CR	-0,8740	0.1465
C3 SR	-2,0509	0.3250
C3 CR	-1,4626	0.3009
C4 SR	-0,6565	0.2481
C4 CR	-0,0682	0.2456
C5 SR	-1,4623	0.1694
C5 CR	-0,8740	0.1465
C6 SR	-2,1590	0.3214
C6 CR	-1,5707	0.3118

Tabela 5.12: Logitos estimados pelo modelo reduzido e seus respectivos erros padrão.

indica bom ajuste. Os valores negativos dos logitos estimados (Tabela 12) indicam que o número de frutos sem escurecimento é inferior ao número de frutos com escurecimento em todas as subpopulações (a razão entre as proporções de frutos com e sem escurecimento é menor que um).

Estrutura II e variável resposta ordinal politômica.

Nesse caso temos as mesmas *subpopulações* da seção anterior e um maior número de *categorias de resposta* ($r = 3$), correspondentes aos graus de escurecimento interno alto (A), baixo (B) e nulo (N). Na Tabela 5.10 são apresentados os dados resultantes da classificação dos frutos nessas categorias, em cada tratamento (SR e CR).

Cada linha da tabela 5.13 corresponde a observações do vetor aleatório

$$Y_{ij} = (Y_{ij1}, Y_{ij2}, Y_{ij3})$$

onde Y_{ijk} é o número de frutos na i -ésima categoria de peso ($i = 1, 2, \dots, 6$) e j -ésimo tratamento ($j = 1, 2$) com o grau de escurecimento k ($k = 1, 2, 3$).

SUBPOPULAÇÃO	A	M	B
C1 SR	4	6	30
C1 CR	8	10	22
C2 SR	7	8	25
C2 CR	9	10	21
C3 SR	3	5	32
C3 CR	5	6	29
C4 SR	3	3	34
C4 CR	4	5	31
C5 SR	1	3	36
C5 CR	4	5	31
C6 SR	2	3	35
C6 CR	4	3	33

Tabela 5.13: Número de frutos com grau de escurecimento interno alto (A), médio (M) e baixo (B) em cada categoria de peso (C1, C2, ... , C6) e tratamento (SR, CR).

Dessa forma o vetor aleatório combinado

$$Y = \begin{pmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{62} \end{pmatrix}$$

é modelado pela distribuição *produto de multinomiais* com função de probabilidade

$$P(Y_{111} = y_{111}, \dots, Y_{623} = y_{623}) = \prod_{i=1}^6 \prod_{j=1}^2 \frac{n_{ij}!}{\prod_{k=1}^3 y_{ijk}!} \prod_{k=1}^3 \pi_{ijk}^{y_{ijk}} \quad (5.28)$$

onde y_{ijk} é o valor observado de Y_{ijk} , n_{ij} o tamanho de amostra em cada subpopulação e π_{ijk} a probabilidade de um fruto selecionado aleatoriamente na subpopulação ij apresentar o grau de escurecimento k . Observe que $\sum_{i=1}^6 \sum_{j=1}^2 y_{ijk} = n_{ij}$ e $\sum_{k=1}^3 \pi_{ijk} = 1$ para todo ij .

Seja

$$\pi = (\pi_{111}, \pi_{112}, \pi_{113}, \dots, \pi_{621}, \pi_{622}, \pi_{623})$$

o vetor de parâmetros que indexa a distribuição de \mathbf{Y} .

O vetor 36×1 de proporções amostrais

$$\mathbf{p} = (p_{111}, p_{112}, p_{113}, \dots, p_{621}, p_{622}, p_{623})$$

é um estimador consistente e não viciado de π .

Esperança e covariância de \mathbf{p} . De (4.4) e (4.5), a esperança de \mathbf{p} é

$$\mathbf{E}(\mathbf{p}) = \begin{pmatrix} E(p_{111}) \\ E(p_{112}) \\ \vdots \\ E(p_{622}) \\ E(p_{623}) \end{pmatrix} = \begin{pmatrix} \pi_{111} \\ \pi_{112} \\ \vdots \\ \pi_{622} \\ \pi_{623} \end{pmatrix}$$

e sua matriz de covariâncias é uma matriz bloco diagonal 36×36 da forma

$$\mathbf{V}_{\mathbf{p}}(\pi) = \begin{bmatrix} V_{11} & 0 & \cdots & 0 & 0 \\ 0 & V_{12} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & V_{61} & 0 \\ 0 & 0 & \cdots & 0 & V_{62} \end{bmatrix} \quad (5.29)$$

onde

$$\mathbf{V}_{ij} = \frac{1}{n_{ij}} \begin{bmatrix} \pi_{ij1}(1 - \pi_{ij1}) & -\pi_{ij1}\pi_{ij2} & -\pi_{ij1}\pi_{ij3} \\ -\pi_{ij1}\pi_{ij2} & \pi_{ij2}(1 - \pi_{ij2}) & -\pi_{ij2}\pi_{ij3} \\ -\pi_{ij1}\pi_{ij3} & -\pi_{ij2}\pi_{ij3} & \pi_{ij3}(1 - \pi_{ij3}) \end{bmatrix} \quad (5.30)$$

corresponde a matriz de covariância do subvetor de \mathbf{p} correspondente à subpopulação ij

e $\mathbf{0}$ é uma matriz 3×3 de zeros.

Escolha da função de resposta. Suponha que o objetivo da pesquisa seja avaliar o efeito dos tratamentos (SR e CR) sobre a percentagem média de área da seção longitudinal do fruto afetada pelo escurecimento interno, em diferentes categorias de peso, mas, por motivos de ordem prática, a avaliação do grau de escurecimento dos frutos foi apenas visual, com posterior classificação nas categorias A, B e N (a área afetada em cada fruto não foi mensurada). A variável peso do fruto (discretizada em seis categorias) foi usada como co-fator.

Deseja-se definir uma função de resposta que represente a percentagem média de área da seção longitudinal afetada pelo escurecimento em cada combinação tratamento \times categoria de peso (subpopulação). Sejam os valores 0,75, 0,25 e 0 que buscam representar a percentagem de área afetada em cada fruto classificado como A, B e N, respectivamente. Assim, a expressão

$$\frac{0,75Y_{ij1} + 0,25Y_{ij2} + 0,00Y_{ij3}}{n_{ij}} = 0,75p_{ij1} + 0,25p_{ij2}$$

denominada *score médio* representa a resposta de interesse.

Definimos então a função de resposta

$$\mathbf{F} : IR^{36} \longrightarrow IR^{12}$$

como

$$\mathbf{F}(\mathbf{p}) = \mathbf{A}\mathbf{p}$$

onde \mathbf{A} é uma matriz bloco diagonal 12×36

$$\mathbf{A} = \begin{bmatrix} A_{11} & 0 & \cdots & 0 & 0 \\ 0 & A_{12} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{61} & 0 \\ 0 & 0 & \cdots & 0 & A_{62} \end{bmatrix}$$

com $A_{ij}, 3 \times 1$

$$A_{ij} = \begin{bmatrix} 0,75 & 0,25 & 0 \end{bmatrix}$$

para todo $i = 1, 2, \dots, 6$ e $j = 1, 2$.

Obtém-se, então

$$\mathbf{F}(\mathbf{p}) = \begin{pmatrix} F_{11}(\mathbf{p}) \\ F_{12}(\mathbf{p}) \\ F_{21}(\mathbf{p}) \\ \vdots \\ F_{61}(\mathbf{p}) \\ F_{62}(\mathbf{p}) \end{pmatrix} = \begin{pmatrix} 0,75p_{111} + 0,25p_{112} \\ 0,75p_{121} + 0,25p_{122} \\ 0,75p_{211} + 0,25p_{212} \\ \vdots \\ 0,75p_{611} + 0,25p_{612} \\ 0,75p_{621} + 0,25p_{622} \end{pmatrix} \quad (5.31)$$

onde $F_{ij}(\mathbf{p})$ representa o escore médio para a categoria de peso i e tratamento j .

Distribuição assintótica de $\mathbf{F}(\mathbf{p})$. Para calcular a distribuição assintótica de $\mathbf{F}(\mathbf{p})$ e conseqüentemente sua esperança e matriz de covariâncias assintóticas será usado o método delta multivariado.

De (4.10), o vetor de proporções amostrais \mathbf{p} satisfaz

$$\sqrt{n}(\mathbf{p} - \boldsymbol{\pi}) \xrightarrow{D} N_{36}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{p}}(\boldsymbol{\pi})) \quad (5.32)$$

onde $n = n_{11} = \dots = n_{62}$ é o tamanho de amostra em cada subpopulação e

$$\Sigma_{\mathbf{p}}(\pi) = n \mathbf{V}_{\mathbf{p}}(\pi). \quad (5.33)$$

$\mathbf{V}_{\mathbf{p}}(\pi)$ está definida em (5.29).

Sejam $\mathbf{H} = \frac{\partial \mathbf{F}}{\partial \mathbf{p}}$ e $\mathbf{G} = \frac{\partial^2 \mathbf{F}}{\partial^2 \mathbf{p}}$ matrizes cujos elementos (k, l) são as derivadas de ordem 1 e 2, respectivamente, da k -ésima componente de $\mathbf{F} = (F_{11}(\mathbf{p}), F_{12}(\mathbf{p}), \dots, F_{62}(\mathbf{p})) \equiv (F_1, F_2, \dots, F_k, \dots, F_{12})$, com respeito ao l -ésimo elemento de $\mathbf{p} = (p_{111}, p_{112}, \dots, p_{262}, \dots, p_{263}) = (p_1, p_2, \dots, p_{36})$. Assim, \mathbf{H} é uma matriz bloco diagonal 12×36

$$\mathbf{H} = \frac{\partial \mathbf{F}}{\partial \mathbf{p}} = \begin{bmatrix} D_{11} & 0 & \cdots & 0 & 0 \\ 0 & D_{12} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & D_{61} & 0 \\ 0 & 0 & \cdots & 0 & D_{62} \end{bmatrix} \quad (5.34)$$

onde D_{ij} , para $i = 1, 2$ e $j = 1, 2, \dots, 6$ é da forma

$$D_{ij} = \begin{bmatrix} \frac{\partial F_{ij}^2}{\partial p_{i,j,1}^2} & \frac{\partial F_{ij}^2}{\partial p_{i,j,2}^2} & \frac{\partial F_{ij}^2}{\partial p_{i,j,3}^2} \end{bmatrix} = \mathbf{A}_{ij}$$

e

$$\mathbf{G} = \frac{\partial^2 \mathbf{F}}{\partial^2 \mathbf{p}} = \mathbf{0} \quad (5.35)$$

onde $\mathbf{0}$ é uma matriz de zeros 12×36 .

De (5.34) e (5.35), observa-se que $\mathbf{F}(\mathbf{p})$ possui derivadas contínuas até ordem 2 para todo $\mathbf{p} \in (0, 1)^{36}$, e conseqüentemente numa região contendo π . Satisfeita essa propriedade, por (4.6), $\mathbf{F}(\mathbf{p})$ tem a seguinte expansão quando $\mathbf{p} \rightarrow \pi$:

$$\mathbf{F}(\mathbf{p}) = \mathbf{F}(\pi) + \mathbf{H}'(\mathbf{p} - \pi) + o(\|\mathbf{p} - \pi\|) \quad (5.36)$$

Por (5.32) e (5.36) vale o teorema que resume o método delta multivariado. Então, de

(4.7) a distribuição assintótica de $\mathbf{F}(\mathbf{p})$ é dada por

$$\sqrt{n}[\mathbf{F}(\mathbf{p}) - \mathbf{F}(\boldsymbol{\pi})] \xrightarrow{\mathcal{D}} N_{2r}(\mathbf{0}, \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}(\boldsymbol{\pi})\mathbf{A}'). \quad (5.37)$$

onde $\boldsymbol{\Sigma}_{\mathbf{p}}(\boldsymbol{\pi})$ é a matriz de covariâncias assintótica de \mathbf{p} definida em (5.33).

A propriedade de normalidade assintótica de $\mathbf{F}(\mathbf{p})$ significa que quando os tamanhos de amostra em cada subpopulação (n_{ij}) são grandes, $\mathbf{F}(\mathbf{p})$ tem distribuição aproximada normal multivariada com

$$\mathbf{E}[\mathbf{F}(\mathbf{p})] \simeq \mathbf{F}(\boldsymbol{\pi}) = \mathbf{A}\boldsymbol{\pi} \quad (5.38)$$

e matriz de covariâncias

$$\begin{aligned} \text{Var}[\mathbf{F}(\mathbf{p})] &\simeq n^{-1}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}(\boldsymbol{\pi})\mathbf{A}' = \\ &= n^{-1}\mathbf{A}(n\mathbf{V}_{\mathbf{p}}(\boldsymbol{\pi}))\mathbf{A}' = \mathbf{A}\mathbf{V}_{\mathbf{p}}(\boldsymbol{\pi})\mathbf{A}' \end{aligned} \quad (5.39)$$

Nesse caso, onde \mathbf{F} é uma transformação linear a esperança e matriz de covariâncias de $\mathbf{F}(\mathbf{p})$ são exatas, apenas a distribuição é aproximada, pois

$$\mathbf{E}(\mathbf{F}(\mathbf{p})) = \mathbf{E}(\mathbf{A}\mathbf{p}) = \mathbf{A}\mathbf{E}(\mathbf{p}) = \mathbf{A}\boldsymbol{\pi}$$

e

$$\text{Var}(\mathbf{F}(\mathbf{p})) = \text{Var}(\mathbf{A}\mathbf{p}) = \mathbf{A}\mathbf{V}_{\mathbf{p}}\mathbf{A}',$$

são idênticas às expressões (5.38) e (5.39).

Denotamos a matriz de covariâncias assintótica de $\mathbf{F}(\mathbf{p})$ por

$$\boldsymbol{\Sigma}_{\mathbf{F}(\mathbf{p})}(\boldsymbol{\pi}) = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}(\boldsymbol{\pi})\mathbf{A}'. \quad (5.40)$$

Substituindo π por \mathbf{p} em (5.40) obtém-se um estimador consistente da matriz de covariâncias assintótica de $\mathbf{F}(\mathbf{p})$, necessária para aplicação do método de quadrados mínimos ponderados.

Testes de hipóteses lineares sobre $\mathbf{F}(\pi)$. Hipóteses lineares de interesse podem ser testadas diretamente sobre $\mathbf{F}(\pi)$ sem necessidade de postulação de um modelo. Alternativamente pode-se postular o modelo saturado

$$\mathbf{E}_A[\mathbf{F}(\mathbf{p})] = \mathbf{X}_1\beta_1 \quad (5.41)$$

onde \mathbf{X} é uma matriz identidade 12×12 e as componentes do vetor de parâmetros β_1 correspondem às componentes de $\mathbf{E}_A[\mathbf{F}(\mathbf{p})] = \mathbf{F}(\pi)$, isto é

$$\beta_1 = \mathbf{F}(\pi) = \mathbf{A}\pi = \begin{pmatrix} 0,75p_{111} + 0,25p_{112} \\ 0,75p_{121} + 0,25p_{122} \\ \vdots \\ 0,75p_{161} + 0,25p_{162} \\ 0,75p_{211} + 0,25p_{212} \\ 0,75p_{221} + 0,25p_{222} \\ \vdots \\ 0,75p_{261} + 0,25p_{262} \end{pmatrix} = \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{16} \\ \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{26} \end{pmatrix} \quad (5.42)$$

Suponha que seja de interesse do pesquisador testar as seguintes hipóteses científicas:

1. Os efeitos dos tratamentos sobre o grau de escurecimento interno são iguais em todas as categorias de peso (hipótese de não interação aditiva entre tratamento e peso);
2. Não existe efeito de tratamento sobre o grau de escurecimento interno.

Correspondente à hipótese científica descrita em (1.) pode ser postulada uma hipótese

estatística representada pelo conjunto de restrições lineares

$$H_{10} : \mu_{11} - \mu_{12} = \mu_{21} - \mu_{22} = \dots = \mu_{61} - \mu_{62}.$$

que o vetor de escores médios deve satisfazer. Sua representação matricial é

$$H_{10} : C_1 \mu = 0$$

onde

$$C_1 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Dessa forma, H_{10} pode ser reescrita como um conjunto de restrições lineares sobre o vetor de parâmetros, isto é

$$H_{10} : C_1 \beta = 0$$

Não havendo evidência para rejeitar a hipótese de ausência de interação aditiva, os tratamentos serão comparados considerando as categorias como replicações. caso contrário, a comparação será feita dentro de cada categoria de peso.

Análise dos resultados. Os escores médios observados e seus respectivos erros padrão são apresentados na Tabela 5.14. Conforme mostrado em (5.42), as componentes do vetor de parâmetros do modelo (5.41) coincidem com esses escores. Como trata-se de um modelo saturado, F pertence ao espaço coluna de X_1 , e conseqüentemente, o valor da estatística de Wald para o teste de ajuste desse modelo é zero (Tabela 5.15).

O valor observado da estatística de teste para a hipótese de ausência de interação

SUBPOPULAÇÃO	FUNÇÃO	ERRO PADRÃO
C1 SR	0.5000	0.0306
C1 CR	0.6188	0.0234
C2 SR	0.4875	0.0292
C2 CR	0.5750	0.0237
C3 SR	0.4688	0.0237
C3 CR	0.5375	0.0237
C4 SR	0.5375	0.0313
C4 CR	0.6313	0.0234
C5 SR	0.4875	0.0264
C5 CR	0.5000	0.0293
C6 SR	0.4938	0.0242
C6 CR	0.5375	0.0208

Tabela 5.14: Escores médios observados e seus respectivos erros padrão para cada categoria de peso (C1, C2, ... , C6) e tratamento (SR, CR).

FONTE	G.L.	QUI-QUADRADO	PROB
INTERCEPTO	1	5084.28	0.0000
PESO	5	19.03	0.0019
TRATAMENTO	1	21.27	0.0000
PESO X TRAT	5	4.98	0,4177
RESÍDUO	0	0.00	.

Tabela 5.15: Análise de variância dos escores médios correspondente ao modelo de efeitos principais com interação.

aditiva definida em H_{10} , ($Q_{C1} = 1,14$) não constitui evidência suficiente para rejeitá-la ($p = 0,9502$) (Tabela 5.15).

Foi então proposto o modelo de efeitos principais

$$E_A(F) = X_2\beta_2 \quad (5.43)$$

onde

$$X_2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & -1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

e as componentes de β_2 têm a seguinte interpretação:

β_0 :média geral;

β_1 a β_5 :efeitos diferenciais das categorias de peso C1 a C5, respectivamente.

β_6 :efeito diferencial do tratamento SR.

A análise de variância do modelo (5.43) é apresentada na Tabela 5.16. O valor da estatística de Wald para avaliação do ajuste do modelo é igual ao valor da estatística do teste da hipótese de não interação ($Q_W = Q_C = 1,14$). O efeito de tratamento sobre o escore médio foi estatisticamente significativo ($p = 0.0085$). As estimativas de quadrados mínimos ponderados dos parâmetros do modelo de efeitos principais são apresentadas na

FONTE	G.L.	QUI-QUADRADO	PROB
INTERCEPTO	1	124,86	0,0000
PESO	5	19,16	0.0018
TRATAMENTO	1	6,93	0.0085
RESÍDUO	5	1,14	0,9502

Tabela 5.16: Análise de variância dos escores médios correspondente ao modelo de efeitos principais.

PARÂMETRO	ESTIMATIVA	ERRO PADRÃO	QUI-QUADRADO	PROB
β_0	0.1186	0,0106	124.86	0.0000
β_1	0.0388	0,0254	2.33	0.1266
β_2	-0,0082	0,0230	0.13	0.7200
β_3	-0.0505	0,0182	3.36	0.0056
β_4	0,0878	0,0283	7,67	0,0019
β_5	-0.0266	0,0222	9.61	0.2321
β_6	-0,0269	0,0102	6,93	0,0085

Tabela 5.17: Estimativas de quadrados mínimos ponderados dos parâmetros do modelo de efeitos principais e seus respectivos erros padrão.

Tabela 5.17. Nessa tabela, observa-se que o parâmetros β_2 e β_5 (efeitos diferenciais das categoria de peso 2 e 5, respectivamente) devem ser retirados do modelo e ajustado um modelo reduzido.

Seja o modelo

$$E_A(\mathbf{F}) = \mathbf{X}_3\beta_3 \quad (5.44)$$

onde \mathbf{X}_3 foi obtida retirando-se a terceira e quinta colunas de \mathbf{X}_2 e as componentes do vetor de parâmetros $\beta_3 = (\beta_0, \beta_1, \beta_3, \beta_4, \beta_6)$ têm a mesma interpretação que a do modelo anterior. A análise de variância do modelo definido em (5.44), estimativas dos parâmetros e escores médios estimados para cada subpopulação são apresentados nas tabelas 5.18, 5.19 e 5.20.

FONTE	G.L.	QUI-QUADRADO	PROB
MODELO	4	27,28	0,0000
RESÍDUO	7	2,92	0,8924

Tabela 5.18: Análise de variância dos escores médios correspondente ao modelo reduzido.

PARÂMETRO	ESTIMATIVA	ERRO PADRÃO	QUI-QUADRADO	PROB
β_0	0,1177	0,0106	123,56	0,0000
β_1	0,0281	0,0238	1,39	0,2382
β_3	-0,0546	0,0179	9,32	0,0023
β_4	0,0739	0,0259	8,15	0,0043
β_6	-0,0261	0,0102	6,56	0,0104

Tabela 5.19: Estimativas de quadrados mínimos ponderados dos parâmetros do modelo reduzido e seus respectivos erros padrão.

SUBPOPULAÇÃO	FUNÇÃO	ERRO PADRÃO
C1 SR	0.1197	0.0306
C1 CR	0.1719	0.0234
C2 SR	0.0916	0.0292
C2 CR	0.1438	0.0237
C3 SR	0.0370	0.0237
C3 CR	0.0892	0.0237
C4 SR	0.1655	0.0313
C4 CR	0.2177	0.0234
C5 SR	0.0916	0.0264
C5 CR	0.1438	0.0293
C6 SR	0.0424	0.0242
C6 CR	0.0964	0.0208

Tabela 5.20: Escores médios estimados no modelo reduzido e seus respectivos erros padrão.

A estatística de Wald indica bom ajuste global ($Q_W = 0,8924$) (Tabela 5.18). Novo modelo poderia ser ajustado retirando-se o parâmetro β_1 .

5.2 Exemplo 2. Comparação de testes de vigor para prever emergência em sorgo.

Nesse exemplo são analisados dados sobre testes de vigor em sementes de sorgo (*Sorghum bicolor*), obtidos por Petrini *et al* (1988). O objetivo desse experimento era avaliar a eficiência de testes de vigor para prever a percentagem de emergência das sementes. Os testes foram aplicados em amostras de seis cultivares de sorgo granífero. Foram observados, para cada cultivar, os resultados dos testes de emergência, germinação, tetrazólio, imersão em cloreto de amônia e testes de envelhecimento precoce das sementes. Os testes de envelhecimento, foram realizados combinando dois níveis de temperatura (42 e 45°C) e quatro níveis de tempo (48, 72, 96, e 120h). Para cada teste foi selecionada uma amostra de 200 sementes, conceitualmente representativa de uma amostra aleatória correspondente à subpopulação definida pela cultivar i e teste j . ($i = 1, 2, \dots, 6$ e $j = 1, 2, \dots, 12$).

Nesse exemplo busca-se apenas mostrar a utilização do método de quadrados mínimos ponderados quando a função de resposta de interesse não se enquadra nos casos descritos por Grizzle *et al* (1969) e Forthofer & Koch (1973). A comparação do método de análise dos dados utilizado em Petrini *et al* (1983) e o método de análise aqui proposto não será objeto dessa dissertação.

Na aplicação do método estatístico proposto foram usados os dados dos testes de emergência (testemunha) e de três testes alternativos (germinação, tetrazólio e imersão em cloreto de amônia), em três cultivares (Contigração 621, Agroceres 1011 e Pioneer B-815).

O objetivo do experimento é comparar a eficiência dos testes, portanto é necessário definir alguma forma de mensurá-la. A idéia é usar uma função das observações que rep-

Cultivar	T	Germinacao	TZVigor	Imersao
Contigrao 621	80	89	83	84
Agroceres 1011	67	80	72	70
Pioneer 8311	58	74	64	65

Tabela 5.21: Resultados dos testes de emergência e envelhecimento precoce (EP) em três cultivares de sorgo granífero.

representa a proximidade entre o resultado obtido pelo teste de emergência (proporção de sementes que emergiram após a germinação) e o obtido por cada um dos testes alternativos. Nos testes de germinação e imersão em cloreto de amônia observa-se a percentagem de sementes germinadas, enquanto no tetrazólio observa-se a percentagem de sementes com embrião uniformemente colorido e extremidades do escutelo não coloridas. Por simplicidade, o resultado em todos os métodos alternativos será referido como percentagem de germinação.

A função de interesse, nesse caso, pode ser arranjada como um vetor (9×1) , $\mathbf{F} = (F_{12}, \dots, F_{14}, \dots, F_{32}, \dots, F_{34})$, onde cada componente representa a eficiência do teste alternativo j na cultivar i , para $i = 1, 2, 3$ e $j = 2, 3, 4$.

Geralmente, para explicar a variabilidade entre as componentes da função \mathbf{F} , busca-se ajustar um modelo linear usualmente expresso por

$$E_A(\mathbf{F}) = \mathbf{X}\beta \quad (5.45)$$

onde \mathbf{X} é uma matriz $9 \times t$ de constantes conhecidas de posto t , β um vetor $t \times 1$ de parâmetros desconhecidos do modelo e $E_A(\cdot)$ denota a esperança assintótica de \mathbf{F} , isto é, quando o número de observações usadas para o cálculo das F_{ij} é grande.

A estrutura dos dados objeto de análise nesse trabalho pode ser resumida na Tabela 5.20.

Outra forma de representar esses dados é através do vetor $\mathbf{Y} = (Y_{11}, \dots, Y_{14}, \dots, Y_{31}, \dots, Y_{34})$, onde Y_{ij} é o número de sementes da cultivar i que germinaram ou emergiram quando

submetidas ao tratamento (método) j .

As variáveis indicadoras de sucesso em cada semente são consideradas mutuamente independentes e as probabilidades de sucesso (germinar ou emergir) são supostas iguais, dentro da mesma subpopulação. Sejam X_l e X_m variáveis indicadoras de sucesso nas l e m -ésima sementes (pertencentes à mesma subpopulação), respectivamente ($l = 1, \dots, 200$). Então a probabilidade de a l -ésima semente apresentar a resposta k_1 dado que a m -ésima semente apresentou a resposta k_2 não depende da distribuição de X_m para todo $m \neq l$, isto é:

$$P(X_l = k_1 | X_m = k_2) = P(X_l = k_1).$$

Seja Y_{ij} o número de sucessos que na cultivar i e tratamento j . Então

$$Y_{ij} = \sum_{m=1}^{200} X_{ijm}$$

onde X_{ijm} é o indicador de sucesso da m -ésima semente na subpopulação ij ($i = 1, 2, 3$ e $j = 1, 2, \dots, 4$). Assim, Y_{ij} é modelado pela distribuição binomial com parâmetros $n = 200$ e π_{ij} , onde π_{ij} representa a probabilidade de uma semente da cultivar i germinar quando submetida ao tratamento j . Como cada y_{ij} provém de amostras independentes, o vetor aleatório \mathbf{Y} tem distribuição produto de binomiais cuja função de probabilidade conjunta é:

$$P(Y_{11} = y_{11}, Y_{12} = y_{12}, \dots, Y_{34} = y_{34}) = \prod_{i=1}^3 \prod_{j=1}^4 \binom{n}{y_{ij}} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{n-y_{ij}} \quad (5.46)$$

onde $\pi_{ij} \in (0, 1)$ para todo $i=1,2,3$ e $j=1,2,\dots,4$.

O vetor de proporções amostrais de sucessos (sementes germinadas ou plântulas que emergiram, dependendo do método) $\mathbf{p} = (p_{11}, \dots, p_{14}, \dots, p_{31}, \dots, p_{34})$, onde $p_{ij} = \frac{Y_{ij}}{n}$ é um estimador consistente (Teorema B1) e não viciado de π .

5.2.1 Esperança e covariância de \mathbf{p} .

O vetor de esperanças e a matriz de covariâncias de \mathbf{p} são expressos por:

$$E(\mathbf{p}) = E \begin{pmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{34} \end{pmatrix} = E \begin{pmatrix} \frac{y_{11}}{n} \\ \frac{y_{12}}{n} \\ \vdots \\ \frac{y_{34}}{n} \end{pmatrix} = \begin{pmatrix} \pi_{11} \\ \pi_{12} \\ \vdots \\ \pi_{34} \end{pmatrix} \quad (5.47)$$

$$\mathbf{V}_{\mathbf{p}}(\pi) = \frac{1}{n} \begin{bmatrix} \pi_{11}(1 - \pi_{11}) & 0 & \dots & 0 \\ 0 & \pi_{12}(1 - \pi_{12}) & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & \pi_{34}(1 - \pi_{34}) \end{bmatrix} \quad (5.48)$$

Substituindo π_{ij} por p_{ij} em [5.47] e [5.48] obtemos estimadores da esperança e matriz de covariâncias de π , respectivamente. As componentes da matriz de covariâncias de \mathbf{p} , denotada por $\mathbf{V}_{\mathbf{p}}(\pi)$, são funções contínuas de π , assim $\mathbf{V}_{\mathbf{p}}(\mathbf{p}) \equiv \mathbf{V}_{\mathbf{p}}$, é também um estimador consistente de $\mathbf{V}_{\mathbf{p}}(\pi)$.

5.2.2 Escolha da função de resposta.

A função \mathbf{F} do vetor de proporções amostrais \mathbf{p} deve ser tal que suas componentes sejam uma representação matemática da idéia de eficiência do método alternativo j na cultivar i ($i = 1, 2, 3$ e $j = 2, 3, 4$). Segundo Koch *et al* (1984), para aplicação do método de *quadrados mínimos ponderados* na estimação de parâmetros do modelo descrito em [5.45], é necessário dispor de um estimador consistente da matriz de covariâncias assintótica de \mathbf{F} . Além disso, a função \mathbf{F} deve ser definida de modo que sua matriz de covariâncias seja não singular.

O método geralmente utilizado para obter a distribuição assintótica de \mathbf{F} e conseqüentemente sua matriz de covariâncias assintótica, é o método delta multivariado. Para aplicação desse método \mathbf{F} deve ser expressa como função de uma estatística assintoticamente normal e satisfazer às condições (i) e (ii) descritas no Apêndice B.

De (4.10), o estimador \mathbf{p} de π na distribuição produto de binomiais satisfaz

$$\sqrt{n}(\mathbf{p} - \pi) \xrightarrow{D} N_{12}(\mathbf{0}, \Sigma_{\mathbf{p}}(\pi)) \quad (5.49)$$

onde n é o tamanho de amostra em cada subpopulação ($n_{ij} = n$ para qualquer i e j) e $\Sigma_{\mathbf{p}}(\pi)$ é a matriz de covariâncias assintótica de \mathbf{p} .

Dado o vetor de proporções amostrais $\mathbf{p} = (p_{11}, \dots, p_{14}, p_{21}, \dots, p_{24}, p_{31}, \dots, p_{34})$, definiremos *distância relativa* (d_{ij}) entre o método alternativo j ($j = 2, 3, 4$) e o método testemunha ($j = 1$) na cultivar i como a razão entre a diferença dos resultados obtidos nos dois métodos (percentagens de emergência e germinação, respectivamente) e o resultado obtido no método testemunha, isto é:

$$d_{ij}(\mathbf{p}) = \frac{p_{i1} - p_{ij}}{p_{i1}}$$

Então, o vetor de distâncias $\mathbf{d}(\mathbf{p})$ constituído pelas componentes $d_{ij}(\mathbf{p})$ é um vetor 9×1 da seguinte forma:

$$\mathbf{d}(\mathbf{p}) = \begin{pmatrix} d_{12}(\mathbf{p}) \\ d_{13}(\mathbf{p}) \\ \cdot \\ \cdot \\ \cdot \\ d_{33}(\mathbf{p}) \\ d_{34}(\mathbf{p}) \end{pmatrix} = \begin{pmatrix} \frac{p_{11} - p_{12}}{p_{11}} \\ \frac{p_{11} - p_{13}}{p_{11}} \\ \cdot \\ \cdot \\ \cdot \\ \frac{p_{31} - p_{33}}{p_{31}} \\ \frac{p_{31} - p_{34}}{p_{31}} \end{pmatrix}$$

Como a idéia de eficiência está relacionada com a distância entre os resultados, não

importando se o método subestima ou superestima a percentagem de emergência, \mathbf{F} deve ser uma função par de d , isto é

$$\mathbf{F}(d) = \mathbf{F}(-d)$$

para todo $d \in \mathbb{R}^9$.

A função que representa a eficiência deve ser decrescente com d , pois quanto maior a distância relativa entre os métodos, menos eficiente ele será. Assim, \mathbf{F} poderia ser definida como

$$\mathbf{F}_1(d) = \mathbf{1} - |d|$$

onde $\mathbf{1}$ é um vetor 9×1 de 1's.

Nesse caso, o método delta não pode ser aplicado para cálculo da distribuição assintótica de \mathbf{F} , pois $\mathbf{F}(\cdot)$ não é diferenciável em $d = 0$ (a condição i não é satisfeita). Entre outras alternativas, as funções

$$\mathbf{F}_2(d) = \mathbf{1} - d^2$$

e

$$\mathbf{F}_3(d) = \exp(-5(d^2))$$

são funções pares e decrescentes de d .

\mathbf{F}_3 apresenta a vantagem de suas componentes estarem restritas ao intervalo $(0, 1)$, o que é desejável numa medida de eficiência. Espera-se, pela natureza do problema que os valores das componentes de \mathbf{F} concentrem-se entre 0 e 0.30. É de interesse, portanto, que a derivada primeira de função que representa a eficiência assumam valores próximos da unidade nessa região, o que significa que iguais variações em d correspondem a iguais variações em \mathbf{F} . Graficamente, isso equivale ao paralelismo dessa função e \mathbf{F}_1 . Na figura 1, observa-se que \mathbf{F}_3 tem melhor comportamento que \mathbf{F}_2 com relação a esse aspecto.

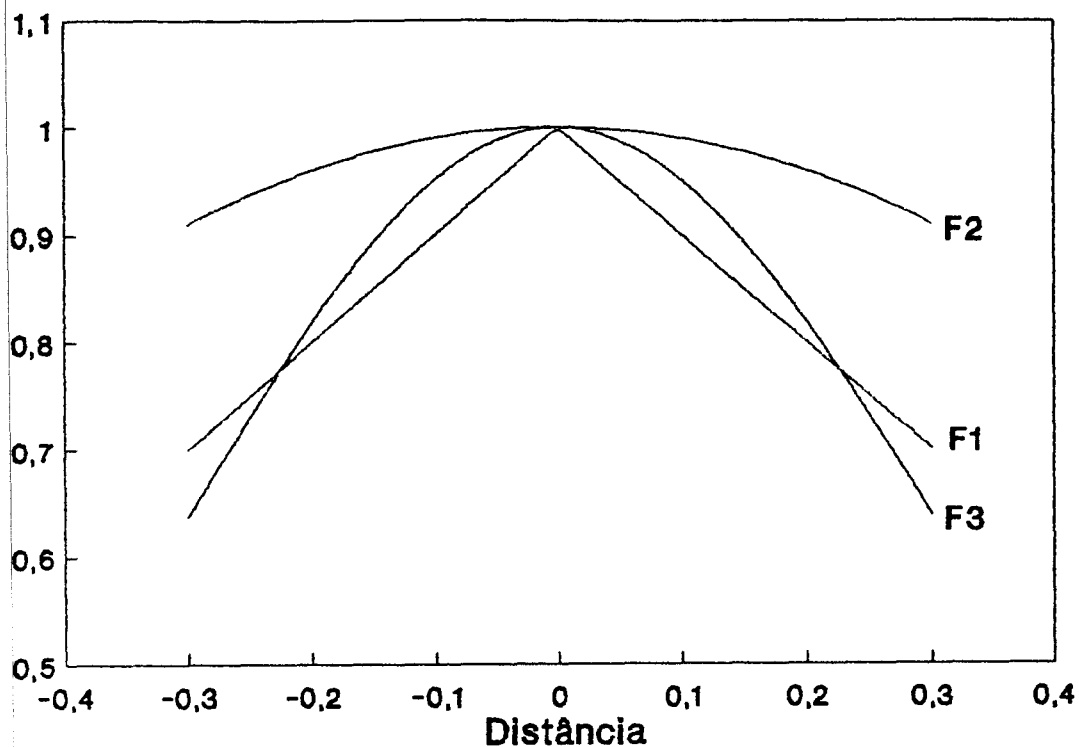


Figura 1. Representação gráfica das funções de resposta F_1 , F_2 e F_3 .

A escolha da função ideal para representar eficiência não será objeto dessa análise. Para tanto, seria de fundamental importância, a participação do pesquisador da área de conhecimento subjacente. Nesse exemplo busca-se apenas mostrar o uso do método de quadrados mínimos ponderados quando a função de interesse não se enquadra nos casos anteriormente abordados (funções lineares e funções logarítmicas).

A função do vetor de proporções amostrais que busca representar a eficiência dos métodos alternativos será então definida como

$$F : R^{12} \mapsto R^9$$

$$\mathbf{F}(\mathbf{p}) = \begin{pmatrix} F_{12}(\mathbf{p}) \\ F_{13}(\mathbf{p}) \\ \vdots \\ F_{33}(\mathbf{p}) \\ F_{34}(\mathbf{p}) \end{pmatrix} = \begin{pmatrix} \exp \left[-5 \left(\frac{p_{11} - p_{12}}{p_{11}} \right)^2 \right] \\ \exp \left[-5 \left(\frac{p_{11} - p_{13}}{p_{11}} \right)^2 \right] \\ \vdots \\ \exp \left[-5 \left(\frac{p_{11} - p_{13}}{p_{11}} \right)^2 \right] \\ \exp \left[-5 \left(\frac{p_{31} - p_{34}}{p_{31}} \right)^2 \right] \end{pmatrix} \quad (5.50)$$

onde cada F_{ij} representa a eficiência do método j na cultivar i .

A verificação das condições i) e ii) para a função definida em [5.50] será feita quando da aplicação do método delta para cálculo de distribuições assintóticas.

5.2.3 Distribuição assintótica de $\mathbf{F}(\mathbf{p})$.

Para calcular a distribuição assintótica de $\mathbf{F}(\mathbf{p})$ e conseqüentemente sua esperança e matriz de covariâncias assintóticas será usado o método delta multivariado descrito no Apêndice B.

Sejam $\mathbf{H} = \frac{\partial \mathbf{F}}{\partial \mathbf{p}}$ e $\mathbf{G} = \frac{\partial^2 \mathbf{F}}{\partial^2 \mathbf{p}}$ matrizes 9×12 cujos elementos (k, l) são as derivadas de ordem 1 e 2, respectivamente, da k -ésima componente de $\mathbf{F} = (F_{12}(\mathbf{p}), F_{13}(\mathbf{p}), \dots, F_{34}(\mathbf{p})) \equiv (F_1, F_2, \dots, F_k, \dots, F_9)$, com respeito ao l -ésimo elemento de $\mathbf{p} = (p_{11}, p_{12}, \dots, p_{31}, \dots, p_{34}) = (p_1, p_2, \dots, p_{12})$. Assim, $\frac{\partial \mathbf{F}}{\partial \mathbf{p}}$ é uma matriz bloco diagonal

$$\mathbf{H} = \frac{\partial \mathbf{F}}{\partial \mathbf{p}} = \begin{bmatrix} \mathbf{D}_1 & 0 & 0 \\ 0 & \mathbf{D}_2 & 0 \\ 0 & 0 & \mathbf{D}_3 \end{bmatrix}$$

onde \mathbf{D}_i , para $i = 1, 2, 3$ é da forma:

$$D_i = \begin{bmatrix} \frac{\partial F_{i2}}{\partial p_{i1}} & \frac{\partial F_{i2}}{\partial p_{i2}} & \frac{\partial F_{i2}}{\partial p_{i3}} & \frac{\partial F_{i2}}{\partial p_{i4}} \\ \frac{\partial F_{i3}}{\partial p_{i1}} & \frac{\partial F_{i3}}{\partial p_{i2}} & \frac{\partial F_{i3}}{\partial p_{i3}} & \frac{\partial F_{i3}}{\partial p_{i4}} \\ \frac{\partial F_{i4}}{\partial p_{i1}} & \frac{\partial F_{i4}}{\partial p_{i2}} & \frac{\partial F_{i4}}{\partial p_{i3}} & \frac{\partial F_{i4}}{\partial p_{i4}} \end{bmatrix} \quad (5.51)$$

onde

$$\begin{cases} \frac{\partial F_{ij}}{\partial p_{kl}} = -10 \left(\frac{p_{i1} - p_{i2}}{p_{i1}} \right) \left(1 + \frac{p_{i2}}{p_{i1}} \right) e^{-5 \left(\frac{p_{i1} - p_{i2}}{p_{i1}} \right)^2}, & \text{se } k = j \\ 0, & \text{se } k \neq j \end{cases}$$

e

$$\begin{cases} \frac{\partial F_{ij}}{\partial p_{kl}} = -10 \left(\frac{p_{i1} - p_{i2}}{p_{i1}} \right) \left(1 - \frac{1}{p_{i1}} \right) e^{-5 \left(\frac{p_{i1} - p_{i2}}{p_{i1}} \right)^2}, & \text{se } k = i \text{ e } j = l \\ 0, & \text{c.c.} \end{cases}$$

De modo semelhante, a matriz de derivadas parciais de segunda ordem, $\frac{\partial^2 F}{\partial^2 p}$, é expressa por

$$G = \frac{\partial F}{\partial p} = \begin{bmatrix} G_1 & & 0 \\ & G_2 & \\ 0 & & G_3 \end{bmatrix}$$

onde G_i , para $i = 1, 2, 3$ é:

$$G_i = \begin{bmatrix} \frac{\partial F_{i2}^2}{\partial p_{i1}^2} & \frac{\partial F_{i2}^2}{\partial p_{i2}^2} & \frac{\partial F_{i2}^2}{\partial p_{i3}^2} & \frac{\partial F_{i2}^2}{\partial p_{i4}^2} \\ \frac{\partial F_{i3}^2}{\partial p_{i1}^2} & \frac{\partial F_{i3}^2}{\partial p_{i2}^2} & \frac{\partial F_{i3}^2}{\partial p_{i3}^2} & \frac{\partial F_{i3}^2}{\partial p_{i4}^2} \\ \frac{\partial F_{i4}^2}{\partial p_{i1}^2} & \frac{\partial F_{i4}^2}{\partial p_{i2}^2} & \frac{\partial F_{i4}^2}{\partial p_{i3}^2} & \frac{\partial F_{i4}^2}{\partial p_{i4}^2} \end{bmatrix} \quad (5.52)$$

De [5.51] e [5.52] observa-se que $F(p)$ possui derivadas contínuas até ordem 2 para todo $p \in (0, 1)^{24}$, e conseqüentemente numa região contendo π . Satisfeita essa propriedade

$\mathbf{F}(\mathbf{p})$ tem a seguinte expansão quando $\mathbf{p} \rightarrow \pi$:

$$\mathbf{F}(\mathbf{p}) = \mathbf{F}(\pi) + \mathbf{H}'(\mathbf{p} - \pi) + o(\|\mathbf{p} - \pi\|) \quad (5.53)$$

Por [5.49] e [5.53] e aplicando o teorema que resume o método delta multivariado, a distribuição assintótica de $\mathbf{F}(\mathbf{p})$ é dada por

$$\sqrt{n}[\mathbf{F}(\mathbf{p}) - \mathbf{F}(\pi)] \xrightarrow{D} N_{12}(\mathbf{0}, \mathbf{H}\Sigma_{\mathbf{p}}(\pi)\mathbf{H}')$$

ou seja, para n grande, $\mathbf{F}(\mathbf{p})$ tem distribuição aproximada normal multivariada com esperança $\mathbf{F}(\pi)$ e matriz de covariâncias $n^{-1}\mathbf{H}\Sigma_{\mathbf{p}}(\pi)\mathbf{H}'$, onde

$$\Sigma_{\mathbf{p}}(\pi) = n^{-1}V_{\mathbf{p}} = \begin{bmatrix} \pi_{11}(1 - \pi_{11}) & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \pi_{12}(1 - \pi_{12}) & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \pi_{34}(1 - \pi_{34}) \end{bmatrix} \quad (5.54)$$

5.2.4 Análise dos resultados.

A primeira hipótese de interesse a ser investigada é a ausência de interação aditiva entre *cultivar* e *método*. Se não houver evidência para rejeição dessa hipótese, os métodos alternativos podem ser comparados, de um modo global, considerando a eficiência média de cada método nas três cultivares; caso contrário, os métodos devem ser comparados dentro de cada cultivar.

Para investigação da variabilidade entre as componentes da função $\mathbf{F}(\cdot)$ definida em [5.50] e posterior teste de hipóteses de interesse, foi proposto o modelo

$$E_A[F(p)] = X_1\beta_1 \quad (5.55)$$

onde

$$X_1 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -1 & -1 & 0 & 0 & 0 & 0 & -1 & -1 \end{bmatrix}$$

e

$$\beta_1 = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_9 \end{pmatrix}$$

cujas componentes têm a seguinte interpretação:

β_0 : eficiência média;

β_1 : efeito diferencial da cultivar C1;

β_2 : efeito diferencial da cultivar C2;

β_3 : efeito diferencial do método *germinação* dentro da cultivar C1;

β_4 : efeito diferencial do método *TZvigor dentro* da cultivar C1;

β_5 : efeito diferencial do método *germinação* dentro da cultivar C2;

β_6 : efeito diferencial do método *TZvigor dentro* da cultivar C2;

β_7 : efeito diferencial do método *germinação* dentro da cultivar C3;

β_8 : efeito diferencial do método *TZvigor dentro* da cultivar C3.

Os efeitos diferenciais do método *imersão dentro* de cada cultivar são tais que a soma dos efeitos diferenciais dentro de uma mesma cultivar seja zero.

Essa parametrização foi escolhida de modo que os β_i 's são facilmente interpretáveis e as hipóteses de interesse podem ser expressas através de um conjunto de restrições lineares sobre o vetor β_1 . Além disso, a base \mathbf{X} correspondente a essa parametrização, tem posto-coluna completo.

A hipótese de não associação aditiva entre cultivar e método foi postulada através dos seguintes contrastes lineares

$$H_0^{(1)} : \begin{cases} \beta_3 - \beta_5 = 0 \\ \beta_3 - \beta_7 = 0 \\ \beta_4 - \beta_6 = 0 \\ \beta_4 - \beta_8 = 0 \end{cases}$$

Sua representação matricial é

$$H_0^{(1)} : \mathbf{C}_1 \beta_1 = 0 \quad (5.56)$$

onde \mathbf{C}_1 é uma matriz 4×8 de posto 4,

$$\mathbf{C}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}$$

O valor da estatística Q_{C_1} foi 3,88 com um valor p correspondente de 0,4227 (Tabela 5.22). Não existindo evidência suficiente contra a hipótese definida em (5.56), postulou-se o seguinte modelo reduzido

$$\mathbf{E}_A[\mathbf{F}(p)] = \mathbf{X}_2 \beta_2 \quad (5.57)$$

onde

$$X_2 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 0 \\ 1 & -1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

e

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_4 \end{bmatrix}$$

onde β_0, β_1 e β_2 têm a mesma interpretação que no modelo anterior, β_3 e β_4 representam os efeitos diferenciais dos métodos *germinação* e *TZvigor*, respectivamente. Observe que o efeito diferencial do método imersão é igual a $-(\beta_3 + \beta_4)$. A análise de variância do modelo definido em [5.57] é apresentada na Tabela 5.23. O valor da estatística do ajuste do modelo corresponde ao qui-quadrado do resíduo ($Q_W = 3,30$), na Tabela 5.23. O valor p correspondente ($p = 0,5097$) não constitui evidência suficiente para rejeição da hipótese representada pelo modelo [5.57].

FONTE	G.L.	QUI-QUADRADO	PROB
INTERCEPTO	1	23,89	0,0000
CULTIVAR	2	3,91	0.1414
METODO	2	10,40	0.0055
CULT X METODO	4	3,30	0.5097
RESÍDUO	0	0.00	.

Tabela 5.22: Análise de variância da eficiência dos métodos alternativos no modelo saturado.

FONTE	G.L.	QUI-QUADRADO	PROB
INTERCEPTO	1	23,89	0,0000
CULTIVAR	2	3,91	0.1414
METODO	2	10,40	0.0055
RESIDUO	4	3,30	0.5097

Tabela 5.23: Análise de variância da eficiência dos métodos alternativos no modelo de efeitos principais.

5.3 Exemplo 3. Efeito do substrato no enraizamento de alporques de urucuzeiro.

Os dados objeto de análise nesse exemplo foram obtidos de Silva *et al* (1993). Foi realizado um experimento com o objetivo de avaliar o efeito de diferentes substratos no enraizamento de alporques de urucuzeiro (*Bixa orellana* L.). As alporquias foram realizadas em quatorze plantas. Em cada planta foram realizados sete alporques (correspondentes aos sete tratamentos). Os tratamentos (substratos) consistiam de:

1. Areia vermelha (A);
2. Barro (B);
3. Pó de madeira (P);
4. $1/2A+1/2B$;

TRATAMENTO	N ^o . DE RAIZES
AREIA	119
BARRO	84
1/2A+1/2B	83

Tabela 5.24: Número total de raízes por tratamento doze semanas após a alporquia.

5. 1/2A+1/2P;
6. 1/2B+1/2P;
7. 1/3A+1/3B+1/3P.

Após doze semanas de realização das alporquias foram observados os números de raízes em cada alporque.

Na análise dos dados utilizamos apenas os dados correspondentes aos tratamentos 1, 2 e 4. Os números totais de raízes observados nos catorze alporques em cada tratamento são apresentados na Tabela 5.24.

Considera-se que o números de raízes em cada tratatamento são conceitualmente representativos de subpopulações infinitas no sentido de equivalência com uma amostra aleatória estratificada.

Supõe-se que as raízes "nascem " segundo um processo de Poisson não homogêneo no tempo, isto é:

1. Os nascimentos ocorrem sozinhos, e não simultâneamente;
2. Os números de raízes que nascem em intervalos de tempo disjuntos são independentes;
3. A probabilidade do nascimento de k raízes no intervalo de tempo $(s, s + t)$ depende de t .

Como o tempo total de acompanhamento é fixo (12 semanas), o número total de raízes que nascem em cada tratamento (Y_i), nesse período, pode ser modelado pela distribuição de Poisson. Assim a função de probabilidade de Y_i é

$$P(Y_i = y_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}$$

com $y_i \in \{0, 1, 2, \dots\}$ e $\mu \in (0, \infty)$ para todo $i = 1, 2, 3$.

Considerando as subpopulações independentes, o vetor aleatório $\mathbf{Y} = (Y_1, Y_2, Y_3)$ tem distribuição produto de Poisson cuja função de probabilidade é

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) = \prod_{i=1}^3 \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}.$$

5.3.1 Esperança e covariância de \mathbf{Y} .

De [4.12] e [4.13], o vetor de esperanças e a matriz de covariâncias de \mathbf{Y} são

$$E(\mathbf{Y}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} \quad (5.58)$$

e

$$\mathbf{V}_{\mathbf{Y}}(\boldsymbol{\mu}) = \mathbf{D}_{\boldsymbol{\mu}} = \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{bmatrix}. \quad (5.59)$$

respectivamente.

Substituindo $\boldsymbol{\mu}$ por \mathbf{y} em [5.59] obtém-se um estimador consistente da matriz de covariâncias assintótica de \mathbf{Y} .

5.3.2 Escolha da função de resposta

O interesse do estudo é explicar a variabilidade entre os números de raízes por alporque esperados para cada tratamento. Nesse caso, o interesse é modelar a taxa de raízes por alporque denotada por

$$\lambda = (\mathbf{D}_N)^{-1} \mathbf{D}_\mu = \begin{bmatrix} \frac{1}{14} & 0 & 0 \\ 0 & \frac{1}{14} & 0 \\ 0 & 0 & \frac{1}{14} \end{bmatrix} \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \mu_2 & 0 \\ 0 & 0 & \mu_3 \end{bmatrix} = \begin{bmatrix} \frac{\mu_1}{14} & 0 & 0 \\ 0 & \frac{\mu_2}{14} & 0 \\ 0 & 0 & \frac{\mu_3}{14} \end{bmatrix}.$$

De [4.15], um estimador consistente de λ é

$$\hat{\lambda} = (\mathbf{D}_N)^{-1} \mathbf{D}_Y = \begin{bmatrix} \frac{Y_1}{14} & 0 & 0 \\ 0 & \frac{Y_2}{14} & 0 \\ 0 & 0 & \frac{Y_3}{14} \end{bmatrix}$$

Nesse caso a função de resposta cuja esperança assintótica deseja-se modelar é a função identidade, isto é,

$$F(\hat{\lambda}) = \hat{\lambda}.$$

De [4.18], um estimador consistente da matriz de covariâncias assintótica de $\hat{\lambda}$ é

$$\begin{aligned} \mathbf{V}_{\hat{\lambda}} &= (\mathbf{D}_N)^{-1} \mathbf{D}_Y (\mathbf{D}_N)^{-1} = \\ &= \begin{bmatrix} \frac{Y_1}{14^2} & 0 & 0 \\ 0 & \frac{Y_2}{14^2} & 0 \\ 0 & 0 & \frac{Y_3}{14^2} \end{bmatrix}. \end{aligned}$$

5.3.3 Definição do modelo e formulação de hipóteses.

Suponha que seja de interesse testar a hipótese de que o número médio de raízes no tratamento $1/2A+1/2B$ é a média aritmética do número médio de raízes nos tratamentos A e B. Essa hipótese pode ser representada por

$$H1_0 : \lambda_3 = \frac{\lambda_1 + \lambda_2}{2}.$$

Se postularmos o modelo identidade saturado

$$E_A(\hat{\lambda}) = X\beta \quad (5.60)$$

onde X é uma matriz identidade 3×3 , as componentes do vetor de parâmetros β correspondem às componentes de λ e H_0 pode ser reescrita como

$$H1_0 : \beta_3 = \frac{\beta_1 + \beta_2}{2}.$$

Sua representação matricial é

$$H1_0 : C_1\beta = 0$$

onde

$$C_1 = \begin{bmatrix} -1/2 & -1/2 & 1 \end{bmatrix}.$$

Outras hipóteses podem ser postuladas para investigar se a substituição de metade do substrato *areia* por *barro* ou vice-versa, afeta o enraizamento dos alporques. Essas hipóteses podem ser representadas pelas restrições lineares

$$H2_0 : \beta_1 = \beta_3$$

e

TRATAMENTO	N ^o . DE RAÍZES/ALPORQUE	ERRO PADRÃO
AREIA	8,5000	0,2082
BARRO	6,0000	0,1750
1/2A+1/2B	5,9300	0,1162

Tabela 5.25: Número de raízes por alporque em cada tratamento doze semanas após a alporquia.

$$H_{3_0} : \beta_2 = \beta_3,$$

respectivamente. Sua representação matricial é

$$H_{2_0} : C_2\beta$$

onde

$$C_1 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$$

e

$$H_{3_0} : C_3\beta$$

onde

$$C_1 = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}.$$

5.3.4 Análise dos resultados.

Os números médios de raízes por alporque observados em cada tratamento e seus respectivos erro padrão são apresentados na Tabela 5.25.

A análise dos contrastes (Tabela 5.26) indica que o número de raízes no tratamento

CONTRASTE	G.L.	QUI-QUADRADO	PROB
MÉDIA ARITMÉTICA	1	35,82	0,0000
SUBS. 1/2A POR 1/2B	1	89,82	0,0000
SUBS. 1/2B POR 1/2A	1	0,08	0,7722

Tabela 5.26: Graus de liberdade, valores observados das estatísticas de teste e valores de p associados para os contrastes definidos em $H1_0$, $H2_0$ e $H3_0$.

com substrato misto ($1/2A+1/2B$) não corresponde à média de A e B ($p < 0.001$); o número de raízes diminui quando metade do substrato areia é substituída por barro ($p < 0.001$) e não se altera quando metade do substrato barro é substituída por areia ($p = 0,7722$).

Não havendo evidência suficiente para rejeitar a hipótese de igualdade entre os números esperados de raízes por alporque nos tratamentos B e $1/2A+1/2B$, propõe-se um modelo reduzido objetivando explicar a variabilidade entre as componentes de λ . Esse modelo incorpora ao modelo anterior (5.60) a restrição sobre λ correspondente à hipótese $H3_0$. Postula-se então o modelo

$$E_A(\hat{\lambda}) = X_2 \beta_2 \quad (5.61)$$

onde

$$X_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

e as componentes de $\beta_2 = (\beta_1, \beta_2)$ têm a seguinte interpretação:

β_1 = número esperado de raízes por alporque quando é utilizado o substrato areia.

β_2 = número esperado de raízes por alporque quando é utilizado o substrato barro ou metade areia + metade barro.

FONTE	G.L.	QUI-QUADRADO	PROB
MÉDIA DE A	1	1666,00	0,0000
MÉDIA DE B OU C	1	2337,00	0,0000
RESÍDUO	1	0,08	0,7722

Tabela 5.27: Análise de variância do modelo reduzido.

TRATAMENTO	N ^o . DE RAÍZES/ALPORQUE	ERRO PADRÃO
AREIA	8,5000	0.2082
BARRO	5,9641	0.1233
1/2A+1/2B	5,9641	0.1233

Tabela 5.28: Valores preditos dos números de raízes por alporque em cada tratamento e seus respectivos erros padrão.

A análise de variância do número de raízes por alporque correspondente ao modelo postulado em [5.60] é apresentada na Tabela 5.27. Obseva-se que o modelo tem bom ajuste ($p = 0,7722$).

Os valores preditos pelo modelo para os números de raízes por alporque em cada tratamento (substrato) são apresentados na Tabela 5.28.

Capítulo 6

Considerações finais.

Nesta dissertação foi abordada a aplicação do método de *quadrados mínimos ponderados* em casos onde é possível aplicar o *método delta* no cálculo da matriz de covariâncias assintótica da função de resposta. Nesses casos, o tamanho de amostra suficiente para aplicação do método está relacionado da qualidade da aproximação da matriz de covariâncias da função de resposta usando expansão em série de Taylor e da convergência da distribuição da função de resposta à distribuição normal multivariada.

Como objetos de pesquisas futuras estão o uso de outros métodos para cálculo da matriz de covariâncias e sua comparação com o método delta. Além disso, é de interesse a comparação dos métodos tradicionais de análise de experimentos agrícolas com os métodos aqui utilizados.

Apêndice A

Algumas definições e teoremas em álgebra linear.

Nesse apêndice são apresentados definições e teoremas importantes para o entendimento dos modelos lineares segundo a abordagem de coordenadas livres apresentada em Arnold (1981). Nessa abordagem busca-se uma interpretação geométrica das estatísticas de interesse no modelo linear, utilizando, principalmente, os conceitos de projeção de vetores em subespaços, norma de projeções e distância entre vetores e subespaços. Para maiores detalhes e provas dos teoremas ver Hoffman & Kunze (1971), Halmos (1958), Arnold (1981) e Mota (1983).

A.1 Espaços vetoriais, subespaços e bases.

Definição A1. Um *espaço vetorial* consiste do seguinte:

- (1) Um corpo F de escalares;
- (2) Um corpo V de objetos denominados vetores;
- (3) Uma regra ou operação dita adição que associa a cada par de vetores \mathbf{v}_1 e \mathbf{v}_2 em

V um vetor $\mathbf{v}_1 + \mathbf{v}_2$ denominado soma de $\mathbf{v}_1 + \mathbf{v}_2$ tais que:

- (a) A adição é comutativa: $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1$;
- (b) A adição é associativa: $(\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3)$;
- (c) Existe um único vetor $\mathbf{0}$ tal que $\mathbf{v} + \mathbf{0} = \mathbf{v}$ para todo $\mathbf{v} \in V$;

(d) Para cada vetor $v \in V$ existe um único vetor $-v \in V$ tal que

$$v + (-v) = 0;$$

(4) Uma regra ou operação dita multiplicação escalar que associa a cada escalar $c \in F$

um vetor cv tal que:

(a) $1v = v$;

(b) $(c_1c_2v) = c_1(c_2v)$;

(c) $c(v_1 + v_2) = cv_1 + cv_2$;

(d) $(c_1 + c_2)v = c_1v + c_2v$.

Definição A2. Um *produto interno* num espaço vetorial (real ou complexo) é uma função (real ou complexa, respectivamente) de um par de vetores x e y , tal que:

(a) $\langle x, y \rangle = \overline{\langle y, x \rangle}$

(b) $\langle a_1x_1 + a_2x_2, y \rangle = a_1 \langle x_1, y \rangle + a_2 \langle x_2, y \rangle$

(c) $\langle x, x \rangle \geq 0$; $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

No caso de espaços vetoriais sobre o corpo dos reais, a conjugação pode ser ignorada.

Definição A3. Seja E um espaço vetorial com produto interno. O número $\sqrt{\langle x, x \rangle}$ é dito *norma* ou *comprimento* do vetor x e denotado por $\|x\|$. Um espaço vetorial sobre o corpo dos reais, dotado de produto interno é chamado *espaço euclidiano*.

As definições e teoremas apresentados a seguir referem-se a espaços vetoriais de dimensão finita com produto interno, em particular o espaço das n -uplas sobre o corpo dos números reais. Os elementos desse espaço vetorial serão representados como:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

onde $y_i \in R$ para $i = 1, 2, \dots, n$.

Definição A4. Sejam x e y vetores quaisquer do subespaço $V \subset R^n$. x e y são ditos *ortogonais* se e somente se $\langle x, y \rangle = 0$.

Definição A5. Seja S um conjunto de vetores em R^n . S é um *conjunto ortogonal* se e somente se $\langle x, y \rangle = 0$ para quaisquer vetores x e y em S .

Se $\langle x, y \rangle = 0$ e $\|x\| = \|y\| = 1$ para quaisquer vetores x e y em S , o conjunto é dito *ortonormal*.

Definição A6 . Sejam S_1 e S_2 subespaços do espaço vetorial E . S_2 é dito (*espaço*) *complemento ortogonal* de S_1 se e somente se y é ortogonal a S_1 para qualquer $y \in S_2$. O complemento ortogonal de S_1 é denotado por S_1^\perp .

Definição A7 . Sejam $v_1, v_2, \dots, v_p \in R^n$. Então u é uma *combinação linear* dos v_i se existem $a_1, a_2, \dots, a_p \in R$ tais que $u = a_1v_1 + a_2v_2 + \dots + a_pv_p$.

Definição A9. Seja V um conjunto tal que $V \subset R^n$. Então V é um *subespaço* se V é fechado para combinações lineares, isto é, para todo $a_1, a_2, \dots, a_p \in R$, $u \in V$.

Sejam $v_1, v_2, \dots, v_p \in V$, onde V é um subespaço. Os v_i *geram* V se todo $v \in V$ pode ser escrito como uma combinação linear dos v_i .

Os v_i são *linearmente independentes* se $a_1v_1 + a_2v_2 + \dots + a_pv_p = 0 \Rightarrow a_1 = a_2 = \dots = a_p = 0$.

Definição A10 . Seja V um subespaço $v_1, v_2, \dots, v_p \in V$. Os v_i são uma *base* de V de eles são *linearmente independentes* e *geram* V . O número de elementos da base é chamado *dimensão* de V .

Teorema A1. Seja V um subespaço, então:

- V tem uma base ortonormal.
- Quaisquer duas bases de V têm o mesmo número de elementos.
- Se v_1, v_2, \dots, v_p formam uma base de V , então todo $v \in V$ pode ser escrito de maneira única como combinação linear dos v_i .
- Se V tem dimensão p , $v_1, v_2, \dots, v_p \in V$ e os v_i são linearmente independentes, então v_1, v_2, \dots, v_p formam uma base de V .
- Seja X uma matriz $n \times p$ de posto r e sejam

$$V = \{v \in R^n : v = Xa \text{ para algum } a \in R^p\}$$

$$W = \{v \in R^n : X'v = 0\}$$

Então a dimensão de V é r e a dimensão de W é $n - r$.

Definição A11. Uma matriz $n \times p$ é uma *base* de V se as colunas de X constituem uma base de V . X é uma base ortonormal se as colunas de X são ortonormais.

Teorema A2. Seja X uma base do espaço p -dimensional $V \subset R^n$. Então:

- a. X é uma matriz $n \times p$ de posto p e $X'X$ é inversível.
- b. O vetor $v \in V$ se e somente se $v = Xb$ para algum $b \in R^p$. O vetor b é único.

A.2 Transformações lineares.

Definição A12. Sejam E_1 e E_2 espaços vetoriais. Uma *transformação linear* de E_1 em E_2 é uma função T de E_1 em E_2 tal que

$$T(cv_1 + v_2) = cT(v_1) + T(v_2)$$

para quaisquer v_1 e v_2 em E_1 e $c \in R$.

Nas definições e teoremas que se seguem E_1 e E_2 são espaços vetoriais sobre R e T uma transformação linear de E_1 em E_2 .

Teorema A3. Seja $A = \{a_1, a_2, \dots, a_j, \dots, a_n\}$ uma base de E_1 e $B = \{b_1, b_2, \dots, b_i, \dots, b_n\}$ vetores arbitrários em E_2 . Então existe exatamente uma transformação linear tal que

$$T(a_j) = b_i$$

para $i = 1, 2, \dots, n$, e $j = 1, 2, \dots, n$.

Teorema A4. A *imagem* da função T é um subespaço de E_2 . Os vetores $v \in E_1$ tais que $T(v) = 0$ constituem um subespaço de E_1 , dito *núcleo* de T .

Definição A13. O *posto* de uma transformação linear T é a dimensão de sua imagem. A *nulidade* de T é a dimensão do seu núcleo.

Teorema A5. A dimensão de um espaço vetorial E é a soma das dimensões da sua imagem e do seu núcleo:

$$\dim(E) = \text{posto}(T) + \text{nulidade}(T)$$

Definição A14 . Um *operador linear* ou *endomorfismo* sobre um espaço vetorial E é uma transformação linear de E em E .

Teorema A6. Se T é uma transformação bijetora (injetora e sobrejetora) de E_1 em E_2 então a função inversa T^{-1} é uma transformação linear de E_2 em E_1 .

Teorema A7. T é uma transformação *não singular* se e somente se leva todo subconjunto linearmente independente de E_1 sobre um conjunto linearmente independente de E_2 .

Teorema A8. Se a dimensão de E_1 é igual à dimensão de E_2 , as seguintes afirmações são equivalentes:

- (a) T é inversível;
- (b) T é não singular;
- (c) A imagem de T é E_2 ;

(c) Se $\{a_1, a_2, \dots, a_n\}$ é uma base arbitrária de E_1 , $\{T(a_1), T(a_2), \dots, T(a_n)\}$ é uma base arbitrária de E_2 .

Definição A15 . Se E_1 e E_2 são espaços vetoriais sobre o mesmo corpo, uma transformação linear bijetora T de E_1 em E_2 é denominada um *isomorfismo* de E_1 em E_2 .

Sejam $A = \{a_1, a_2, \dots, a_j, \dots, a_n\}$ uma base de E_1 , $B = \{b_1, b_2, \dots, b_i, \dots, b_m\}$ uma base de E_2 e T uma transformação arbitrária E_1 em E_2 . Então cada um dos n vetores $T(a_j)$ pode ser expresso de modo único como uma combinação linear de $b_1, b_2, \dots, b_i, \dots, b_m$

$$T(a_j) = \sum_{i=1}^m x_{ij} b_i$$

ou seja.

$$T(\mathbf{a}_1) = x_{11}\mathbf{b}_1 + x_{21}\mathbf{b}_2 + \dots + x_{m1}\mathbf{b}_m$$

$$T(\mathbf{a}_2) = x_{12}\mathbf{b}_1 + x_{22}\mathbf{b}_2 + \dots + x_{m2}\mathbf{b}_m$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$T(\mathbf{a}_n) = x_{1n}\mathbf{b}_1 + x_{2n}\mathbf{b}_2 + \dots + x_{mn}\mathbf{b}_m$$

Os escalares x_{1j}, \dots, x_{mj} são as coordenadas dos vetores $T(\mathbf{a}_j)$ com relação à base $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ para $j = 1, 2, \dots, n$. Então a transformação T é determinada pelos $m \times n$ escalares x_{ij} dadas as bases \mathbf{A} e \mathbf{B} .

A matriz $m \times n$ definida por $X(i, j) = x_{ij}$ e pela relação

$$T(\mathbf{a}_j) = \sum_{i=1}^m x_{ij} \mathbf{b}_i$$

é denominada *matriz da transformação* T com relação ao par de bases \mathbf{A} e \mathbf{B} .

Teorema A9. Sejam $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ uma base de E_1 . $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ uma base de E_2 . Para cada transformação linear T de E_1 em E_2 , existe uma matriz \mathbf{X} , $m \times n$ que é a matriz de T com relação às bases \mathbf{A} e \mathbf{B} , tal que:

$$[T(\mathbf{v})]_{\mathbf{B}} = \mathbf{X}[\mathbf{v}]_{\mathbf{A}}$$

para todo vetor $\mathbf{v} \in V$.

A.3 Projetores.

Nesta seção serão apresentados teoremas e definições relacionados com uma classe particular de transformações lineares, os *operadores projeção* ou simplesmente *projetores*, com ênfase dos *projetores ortogonais*.

Definição A16. Sejam V e W subespaços de um espaço vetorial E . O espaço E é dito *soma direta* de V e W se e somente se todo vetor $z \in E$ pode ser escrito como

$$z = x + y$$

com $x \in V$ e $y \in W$.

A operação soma direta de dois subespaços é denotada por

$$E = V \oplus W.$$

Se $V \perp W$, a soma direta é dita *ortogonal* e representada por

$$E = V \overset{\perp}{\oplus} W.$$

Nesse caso, o subespaço W é dito *complemento ortogonal* de V e vice-versa.

Definição A17. Sejam V e W subespaços do espaço vetorial E tais que

$$E = V \overset{\perp}{\oplus} W$$

e T um endomorfismo de E . E é um *projetor* ou *operador projeção* de E em V segundo a direção dos vetores de W , se e somente se, para todo $z \in E$ tal que $z = x + y$,

$$Ty = x.$$

Teorema A10. Um endomorfismo T de E é um projetor em algum subespaço se e somente se T é *idempotente*, isto é, $T = T^2$.

Teorema A11. Se P é um projetor de E em V segundo a direção dos vetores de W , então V é a imagem de P e W o núcleo de P e conseqüentemente:

(a) $Py = y$ para todo $y \in V$.

(b) $Py = 0$ para todo $y \in W$.

Teorema A12. Uma transformação linear T é um projetor em algum subespaço se e somente se $(I - T)$ é também um projetor. Se P é um projetor de E em V segundo a direção dos vetores de W , então $(I - P)$ é um projetor de E em W segundo a direção dos vetores de V .

Teorema A13. Sejam P_1 e P_2 projetores do espaço vetorial E . Então:

(a) $Im(P_1) \subset Im(P_2)$ se e somente se $P_2(P_1(y)) = P_1(y)$ para qualquer $y \in E$.

(b) $(P_1 + P_2)$ é um projetor de E se e somente se $P_2(P_1(y)) = P_1(P_2(y)) = 0$, para todo $y \in E$.

(c) $(P_1 - P_2)$ é um projetor de E se e somente se $P_2(P_1(y)) = P_1(P_2(y)) = P_2(y)$, para todo $y \in E$.

Quando um espaço vetorial E é decomposto em dois subespaços ortogonais, isto é,

$$E = V \oplus V^\perp,$$

associada a essa decomposição existe uma classe especial de projetores. os *projetores ortogonais*.

Definição A18. Seja V um subespaço do R^n . O complemento ortogonal de V , denotado por V^\perp é o conjunto de todos os vetores $v \in R^n$ ortogonais a V .

Seja $W \subset V$ um subespaço. Então o conjunto de todos os vetores em V que são ortogonais a W , ou seja $V \cap W^\perp$ será denotado por $V | W$.

Lema 1. V^\perp e $V | W$ são subespaços.

Definição A19. Seja V um subespaço do R^n e W o complemento ortogonal de V , isto é:

$$R^n = V \oplus W.$$

Um projetor P de R^n é um *projetor ortogonal* se e somente se P projeta vetores do R^n em V segundo a direção de W .

Teorema A14. Seja V um subespaço do R^n e $y \in R^n$, então:

(a). Para toda matriz $X_{n \times p}$ de posto $r \leq \min(n, p)$, existe uma matriz $B_{p \times n}$ tal que

$$X'XB = X'$$

XB é a matriz do projetor ortogonal de R^n sobre o $C(X)$;

(b). A projeção é única;

(c). Se a matriz $X_{n \times p}$ de posto p é uma base de V , então $X'X$ é inversível e a projeção de y de sobre V é dada por $X(X'X)^{-1}X'y$.

A projeção de y de sobre o subespaço V será denotada por $P_V y$. P_V é um endomorfismo do R^n (transformação linear de R^n em R^n) que associa a cada vetor $y \in R^n$, sua projeção ortogonal em V . A notação P_V também será usada para a matriz da transformação P_V com relação a uma base qualquer do R^n .

Uma das mais importantes características da matriz do projetor ortogonal (P_V) é que ela independe da base de V escolhida. Essa propriedade é enunciada através do seguinte teorema:

Teorema A15. Sejam X e A duas bases quaisquer do subespaço V (X e A são matrizes $n \times p$ de posto p , tais que $C(X) = C(A)$), então

$$P_V = A(A'A)^{-1}A' = X(X'X)^{-1}X'$$

A seguir serão enunciadas algumas propriedades da matriz P_V que são conseqüências das características dos projetores ortogonais:

(a) $(I - P_V)$ é a matriz do projetor ortogonal sobre V^\perp .

(b) $P_V y$ é a projeção ortogonal de y em V e $(I - P_V)y$ é a projeção ortogonal de y em V^\perp para qualquer vetor $y \in R^n$.

Teorema A16. (Teorema de Pitágoras).

- a. Se $v \in V$, então $\|y - v\|^2 = \|y - P_v y\|^2 + \|P_v y - v\|^2$.
- b. $\|y - P_v y\|^2 \leq \|y - v\|^2$ para todo $v \in V$ valendo a igualdade se e somente se $v = P_v y$.
- c. $P_v y = y$ se e somente se $y \in V$. $P_v y = 0$ se e somente se $y \in V^\perp$.
- d. $P_{v^\perp} y = y - P_v y$. $\|P_{v^\perp} y\|^2 = \|y\|^2 - \|P_v y\|^2$.
- e. Seja $W \subset V$ um subespaço, então $P_v y = P_v(P_w y) = P_w(P_v y)$.
- f. Seja $W \subset V$ um subespaço, então $P_{v|W} y = P_v y - P_w y$. $\|P_{v|W} y\|^2 = \|P_v y\|^2 - \|P_w y\|^2$.
- g. Se $V \perp W$, então $P_v(P_w y) = (P_v P_w) y = 0$.

Teorema A17. Seja X uma base ortonormal do subespaço V . Então:

- (a) $X'X = I$;
- (b) $P_v y = X'Xy$;
- (c) $\|P_v y\|^2 = \|X'Xy\|^2$.

Lema 1. Sejam x_1, x_2, \dots, x_n vetores não nulos mutuamente ortogonais. Então os x_i são linearmente independentes.

Lema 2. Seja $W \subset V \subset R^n$. Então:

- (a) $\dim(W^\perp) = n - \dim(W)$.
- (b) $\dim(V | W) = \dim(V) - \dim(W)$.

Lema 3.

- (a) $(V^\perp)^\perp = V$;
- (b) Seja $W \subset V$. Então $V | (V | W) = W$.

Teorema A18. A é a matriz de um projetor ortogonal se e somente se A é *idempotente e simétrica*, isto é, $A = A^2$ e $A = A'$. Se A é um projetor ortogonal, então $\dim(A) = \text{tr}(A) = \text{posto}(A)$.

Apêndice B

Propriedades assintóticas de estimadores e estatísticas de teste.

Nesse apêndice serão apresentadas definições e teoremas relacionados com *propriedades assintóticas* das estatísticas de interesse na aplicação do método de *quadrados mínimos ponderados*:

- (a) *Estatística de Wald* para o teste do ajuste do modelo (Q_W);
- (b) *Função de resposta estimada* ($F(\mathbf{y})$);
- (c) *Estimador de quadrados mínimos ponderados* do vetor de parâmetros no modelo linear geral (b);
- (d) *Estatísticas de teste* de hipóteses lineares sobre o vetor de parâmetros β .

Para maiores detalhes e provas dos teoremas deste apêndice, ver Bickel & Doksum (1977), Leite & Singer (1980), James (1981), Bishop *et al* (1979) e Arnold (1981).

B.1 Algumas propriedades assintóticas de estimadores.

Seja $\{P_\theta : \theta \in \Theta\}$ uma família de distribuições de um vetor aleatório $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$, indexada pelo parâmetro $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ e $\{\hat{\theta}_n(\mathbf{Y}) = (\hat{\theta}_n^{(1)}(\mathbf{Y}), \hat{\theta}_n^{(2)}(\mathbf{Y}), \dots, \hat{\theta}_n^{(p)}(\mathbf{Y})) : n \geq 1\}$ uma seqüência de estimadores de θ que são similarmente gerados para cada tamanho de amostra.

Definição B1. $\hat{\theta}_n$ é dito um estimador *consistente* de θ se e somente se $\hat{\theta}_n$ converge

em probabilidade para θ , isto é

$$P_{\theta_i}(|\hat{\theta}_n^{(i)} - \theta_i| \geq \epsilon_i) \xrightarrow{n \rightarrow \infty} 0$$

para qualquer $\epsilon_i > 0$, ($i = 1, 2, \dots, p$).

Teorema B1. Seja $\mathbf{Y} = (Y_1, Y_2, \dots, Y_k)$ um vetor aleatório com distribuição multinomial com parâmetros n e $\pi = (\pi_1, \pi_2, \dots, \pi_k)$. Se $\hat{\theta}_n$ é um estimador obtido pelo princípio da substituição de frequências,

$$\hat{\theta}_n(\mathbf{Y}) = \left(\frac{Y_1}{n}, \frac{Y_2}{n}, \dots, \frac{Y_k}{n} \right)$$

então $\hat{\theta}_n$ é um estimador consistente de π .

Teorema B2. Seja $\mathbf{Y} = (Y_1, Y_2, \dots, Y_k)$ um vetor aleatório com distribuição produto de Poisson, com parâmetro $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$. Se $\hat{\theta}_n$ é um estimador obtido pelo princípio da substituição de frequências,

$$\hat{\theta}_n(\mathbf{Y}) = (Y_1, Y_2, \dots, Y_k)$$

então $\hat{\theta}_n$ é um estimador consistente de λ .

Teorema B3. (Teorema central do limite multivariado). Sejam $\mathbf{Y}_1, \mathbf{Y}_2, \dots$ vetores aleatórios k -dimensionais, independentes e identicamente distribuídos. Suponha que \mathbf{Y}_1 tenha variância finita, e sejam μ e Σ a esperança e a matriz de covariâncias de \mathbf{Y}_1 , respectivamente. Seja $\bar{\mathbf{X}}_n$, a média amostral, definida como a média aritmética dos vetores $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$. Então

$$\sqrt{n}(\bar{\mathbf{X}}_n - \mu) \xrightarrow{D} N(\mathbf{0}, \Sigma)$$

quando $n \rightarrow \infty$.

Sejam $\mu_n(\theta)$ e $\Sigma_n(\theta)$ o vetor de esperanças e a matriz de covariâncias de $\hat{\theta}_n$, respectivamente.

Definição B2. Uma seqüência de estimadores $\{\hat{\theta}_n(\mathbf{Y}) : n \geq 1\}$ tem a propriedade de normalidade assintótica se e somente se

$$\sqrt{n}(\hat{\theta}_n - \mu_n(\theta)) \xrightarrow{D} N_k(\mathbf{0}, \Sigma_n(\theta)).$$

Quando essa aproximação é válida, $\mu_n(\theta)$ é dita *esperança assintótica de $\hat{\theta}_n$* e $\Sigma_n(\theta)$ *variância assintótica de $\hat{\theta}_n$* . A expressão

$$(\mathbf{D}_{\Sigma_n(\theta)})^{-1}(\mu_n(\theta) - \theta)$$

onde $\mathbf{D}_{\Sigma_n(\theta)}$ é uma matriz diagonal $p \times p$ cujos elementos da diagonal são as variâncias de $\hat{\theta}_n^{(1)}(\mathbf{Y}), \hat{\theta}_n^{(2)}(\mathbf{Y}), \dots, \hat{\theta}_n^{(p)}(\mathbf{Y})$, respectivamente, é dita *vício padronizado assintótico de $\hat{\theta}_n$* .

Definição B3. Dizemos que um estimador $\hat{\theta}_n$ de θ é *assintoticamente não viciado* quando seu vício padronizado converge a zero quando $n \rightarrow \infty$, isto é,

$$(\mathbf{D}_{\Sigma_n(\theta)})^{-1}(\mu_n(\theta) - \theta) \xrightarrow{n \rightarrow \infty} 0$$

Suponha que $\mathbf{T}^{(1)} = \{\mathbf{T}_n^{(1)} : n \geq 1\}$ e $\mathbf{T}^{(2)} = \{\mathbf{T}_n^{(2)} : n \geq 1\}$ são duas seqüências de estimadores cujas matrizes de covariâncias assintóticas $\Sigma_{n1}(\theta)$ e $\Sigma_{n2}(\theta)$ satisfazem

$$n\Sigma_{ni}(\theta) \xrightarrow{n_i \rightarrow \infty} \Sigma_i(\theta)$$

para $i = 1, 2$ e são seqüências assintoticamente eficientes, isto é,

$$(\mathbf{D}_{\Sigma_{ni}(\theta)})^{-1}(\mu_{ni}(\theta) - \theta) \xrightarrow{n_i \rightarrow \infty} 0$$

Definição B4. Definimos *eficiência assintótica* de $\mathbf{T}^{(1)}$ com relação a $\mathbf{T}^{(2)}$ como a razão entre os determinantes das matrizes de covariância assintótica de $\mathbf{T}^{(1)}$ e $\mathbf{T}^{(2)}$, respectivamente:

$$\xi_A = \frac{|\sum_{n1}(\theta)|}{|\sum_{n2}(\theta)|}$$

Seja um processo de estimação que gera uma seqüência de estimadores $\{\hat{\theta}_n(\mathbf{Y}) : n \geq 1\}$ que é assintoticamente normal e satisfaz

$$n \sum_n(\theta) \xrightarrow{n \rightarrow \infty} \sum(\theta), e$$

$$\sqrt{n}(\mu_n(\theta) - \theta) \xrightarrow{n \rightarrow \infty} \mathbf{0}.$$

Dizemos que a seqüência é assintoticamente eficiente se

$$\sum(\theta) = [\mathbf{I}_\theta]^{-1}.$$

Definição B5. Um processo de estimação que gera uma seqüência de estimadores *assintoticamente normal e assintoticamente eficiente* é dito um *processo de estimação BAN (Best asymptotically normal)*. Os estimadores obtidos por esse processo são ditos *estimadores BAN*

O método de quadrados mínimos ponderados é geralmente utilizado para análise de dados modelados pelas distribuições produto de *Poisson* ou de *multinomiais*. Nessas distribuições os estimadores dos parâmetros que as indexam (π e λ , respectivamente) têm comportamento assintótico que permite o uso do método delta para o cálculo de distribuições assintóticas de funções dessas estimativas que satisfazem às condições (i) e (ii) descritas na seção 3.2.1.

Será apresentado a seguir, o método delta conforme descrito por Bishop *et al* (1979).

B.1.1 O método δ para o cálculo de distribuições assintóticas.

Essa é uma técnica geral importante, aplicada para calcular distribuições assintóticas e conseqüentemente deduzir variâncias e covariâncias assintóticas.

A versão unidimensional do método delta (δ).

Para aplicação do método delta é necessário uma variável aleatória ($\hat{\theta}_n$) cuja distribuição depende de um parâmetro real (θ) de tal forma que:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, \sigma^2(\theta)) \quad (\text{B.1})$$

e uma função $F(x)$ diferenciável em $x = \theta$ que possui a seguinte expansão em torno de θ :

$$F(x) = F(\theta) + F'(\theta)(x - \theta) + o(\|x - \theta\|) \quad (\text{B.2})$$

quando $x \rightarrow \theta$.

O método δ para encontrar médias e variâncias aproximadas é justificado pelo seguinte teorema:

Teorema B4. Se $\hat{\theta}_n$ é uma variável aleatória real e θ um parâmetro real tais que B.1 e B.2 valem então a distribuição assintótica de $F(\hat{\theta}_n)$ é dada por:

$$\sqrt{n} [F(\hat{\theta}_n) - F(\theta)] \xrightarrow{D} N(0, \sigma^2(\theta) [F'(\theta)]^2) \quad (\text{B.3})$$

Esse resultado também pode ser interpretado da seguinte forma: para n grande, $F(\hat{\theta}_n)$ tem distribuição aproximadamente normal com média $F(\theta)$ e variância $n^{-1} \sigma^2(\theta) [f'(\theta)]^2$.

A versão multivariada do método delta (δ).

Seja $\hat{\theta}_n = (\hat{\theta}_{n1}, \hat{\theta}_{n2}, \dots, \hat{\theta}_{np})$ um vetor aleatório p -dimensional e $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ um vetor de parâmetros também p -dimensional. Consideremos que $\hat{\theta}_n$ e θ são tais que:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N_p(0, \Sigma(\theta)) \quad (\text{B.4})$$

onde $\Sigma(\theta)$ é a matriz de covariância assintótica de $\hat{\theta}_n$.

Isso equivale a dizer que para n grande $\hat{\theta}_n$ tem distribuição aproximada normal mul-

tivariada com média θ e matriz de covariâncias $n^{-1} \Sigma(\theta)$.

Agora suponha \mathbf{F} uma função definida num subconjunto aberto do espaço p -dimensional tomando valores no espaço r -dimensional. Consideramos que f é diferenciável em θ , isto é, \mathbf{F} tem a seguinte expansão quando $\mathbf{x} \rightarrow \theta$.

$$F_i(\mathbf{x}) = F_i(\theta) + \sum_{j=1}^p (x_j - \theta_j) \left[\frac{\partial F_i}{\partial x_j} \right]_{\mathbf{x}=\theta} + o(\|\mathbf{x} - \theta\|) \quad (\text{B.5})$$

para $i = 1, 2, \dots, p$.

Seja $\left[\frac{\partial \mathbf{f}}{\partial \theta} \right]$ uma matriz $r \times p$ cujo elemento (i,j) é a derivada parcial de F_i com respeito à j -ésima coordenada de $\mathbf{x} = (x_1, x_2, \dots, x_p)$, avaliada em $\mathbf{x} = \theta$, isto é:

$$\left(\frac{\partial \mathbf{F}}{\partial \theta} \right)_{ij} = \left[\frac{\partial F_i}{\partial x_j} \right]_{\mathbf{x}=\theta}$$

então a expressão em B.5 pode ser resumida como:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\theta) + \left(\frac{\partial \mathbf{F}}{\partial \theta} \right) (\mathbf{x} - \theta) + o(\|\mathbf{x} - \theta\|) \quad (\text{B.6})$$

quando $\mathbf{x} \rightarrow \theta$.

Considerando as notações acima, o método δ multivariado, pode ser estabelecido através do seguinte teorema:

Teorema B5. Sejam $\hat{\theta}_n$, θ e \mathbf{F} conforme descritos anteriormente e suponha que valem B.1 e B.6. Então a distribuição assintótica de $\mathbf{F}(\hat{\theta}_n)$ é dada por:

$$\sqrt{n}(\mathbf{F}(\hat{\theta}_n) - \mathbf{F}(\theta)) \xrightarrow{D} N_p(0, \left(\frac{\partial \mathbf{F}}{\partial \theta} \right) \Sigma(\theta) \left(\frac{\partial \mathbf{F}}{\partial \theta} \right)') \quad (\text{B.7})$$

Agora suponha que $\theta = \theta(\varphi)$, onde φ é um vetor s -dimensional de parâmetros. Considere também que a matriz de covariâncias assintótica em B.1 depende de φ , isto é:

$$\sqrt{n}(\hat{\theta}_n - \theta(\varphi)) \xrightarrow{D} N_p(0, \Sigma(\varphi)) \quad (\text{B.8})$$

Finalmente, seja $\left(\frac{\partial \mathbf{F}}{\partial \theta} \right)$, a matriz cujo elemento (i,j) é a derivada parcial de F_i com

respeito ao j -ésimo elemento de $\mathbf{x} = (x_1, x_2, \dots, x_p)$, avaliada em $\mathbf{x} = \theta(\varphi)$. Então a distribuição assintótica de $\mathbf{F}(\hat{\theta}_n)$ é dada por:

$$\sqrt{n}[\mathbf{F}(\hat{\theta}_n) - \mathbf{F}(\theta(\varphi))] \xrightarrow{D} N_p(0, \left(\frac{\partial \mathbf{F}}{\partial \theta}\right) \Sigma(\theta(\varphi)) \left(\frac{\partial \mathbf{F}}{\partial \theta}\right)') \quad (\text{B.9})$$

B.2 Estatísticas de teste com distribuição assintótica χ^2 .

Os teoremas apresentados a seguir são úteis no cálculo das distribuições assintóticas de Q_W e Q_C .

Teorema B6. Sejam $\mathbf{F}(\hat{\theta}_n)$ e $\mathbf{F}(\theta)$ são tais que a condição [B.7] é satisfeita. Se \mathbf{A} é uma matriz $m \times p$ qualquer, então

$$\sqrt{n}[\mathbf{A}\mathbf{F}(\hat{\theta}_n) - \mathbf{A}\mathbf{F}(\theta)] \xrightarrow{D} N_m(0, \mathbf{A}\Sigma_{\mathbf{F}}(\theta)\mathbf{A}')$$

onde $\Sigma_{\mathbf{F}}(\theta)$ é a matriz de covariâncias assintótica de $\mathbf{F}(\hat{\theta}_n)$.

O teorema acima é um caso particular de aplicação do método delta multivariado (Teorema B5).

Teorema B7. Seja \mathbf{Y} um vetor aleatório u -dimensional com distribuição normal multivariada com esperança μ e matriz de covariâncias Σ . Então a forma quadrática

$$\mathbf{Y}'\Sigma\mathbf{Y}$$

tem distribuição χ^2 com u graus de liberdade.

Teorema B8. Seja $\{\hat{\theta}_n(\mathbf{Y}) = (\hat{\theta}_n^{(1)}(\mathbf{Y}), \hat{\theta}_n^{(2)}(\mathbf{Y}), \dots, \hat{\theta}_n^{(u)}(\mathbf{Y})) : n \geq 1\}$ uma seqüência de estimadores tal que

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N_u(0, \Sigma(\theta)).$$

Então a seqüência de formas quadráticas

$$(\hat{\theta}_n)'[\Sigma(\hat{\theta}_n)](\hat{\theta}_n),$$

onde $\Sigma(\hat{\theta}_n)$ é um estimador consistente da matriz de covariâncias assintótica de $\hat{\theta}_n$, tem distribuição assintótica χ^2 com u graus de liberdade.

Apêndice C

Métodos computacionais.

Para análise dos exemplos apresentados no capítulo 5, foram utilizados o módulo IML (Interactive matrix analysis) e o procedimento CATMOD (Categorical data analysis) do módulo STAT (Statistics) do sistema SAS. O módulo IML foi usado para cálculo da matriz de covariâncias assintótica da função de resposta nos exemplos 2 e 3. O procedimento CATMOD calcula essa matriz internamente quando os dados objeto de análise são modelados pela distribuição produto de multinomiais e a função de resposta resulta de combinações de funções lineares, exponenciais ou logarítmicas. No exemplo 2, a função de resposta não pertence a essa classe e no exemplo 3 os dados são modelados pela distribuição produto de Poisson.

Neste apêndice constam as listagens dos programas utilizados nos seguintes casos:

1. Exemplo 1: estrutura I - resposta dicotômica.
2. Exemplo 1: estrutura II - resposta dicotômica.
3. Exemplo 1: estrutura II - resposta politômica.
4. Exemplo 2 : modelo saturado.
5. Exemplo 2: modelo reduzido.
6. Exemplo 3.

```

0 0 0 0 0 0,
ALL_PARMS 1 0 0 -1 0 0
0 0 0 0 0 0,
ALL_PARMS 1 0 0 0 -1 0
0 0 0 0 0 0,
ALL_PARMS 1 0 0 0 0 -1
0 0 0 0 0 0;

```

```

CONTRAST 'NÃO ASS.PESO X ESC EM CR' ALL_PARMS 0 0 0 0 0 0
1 -1 0 0 0 0,
ALL_PARMS 0 0 0 0 0 0
1 0 -1 0 0 0,
ALL_PARMS 0 0 0 0 0 0
1 0 0 -1 0 0,
ALL_PARMS 0 0 0 0 0 0
1 0 0 0 -1 0,
ALL_PARMS 0 0 0 0 0 0
1 0 0 0 0 -1;

```

```

CONTRAST 'HOMOGENEIDADE DA ASSOCIAÇÃO' ALL_PARMS 1 -1 0 0 0 0
-1 1 0 0 0 0,
ALL_PARMS 1 0 -1 0 0 0
-1 0 1 0 0 0,
ALL_PARMS 1 0 0 -1 0 0
-1 0 0 1 0 0,
ALL_PARMS 1 0 0 0 -1 0
-1 0 0 0 1 0,
ALL_PARMS 1 0 0 0 0 -1
-1 0 0 0 0 1;

```

RUN;

TITLE 'ESTRUTURA II - RESPOSTA DICOTÔMICA';

DATA ABCX21;

INPUT PESO \$ TRAT \$ ESC \$ CONT @@;

CARDS;

C1 SR S 10 C3 SR S 8 C5 SR S 4

C1 SR N 30 C3 SR N 32 C5 SR N 36

C1 CR S 18 C3 CR S 11 C5 CR S 9

C1 CR N 22 C3 CR N 29 C5 CR N 31

C2 SR S 15 C4 SR S 6 C6 SR S 5

C2 SR N 25 C4 SR N 34 C6 SR N 35

C2 CR S 19 C4 CR S 9 C6 CR S 7

C2 CR N 21 C4 CR N 31 C6 CR N 33

;
PROC CATMOD ORDER=DATA;

WEIGHT CONT;

POPULATION PESO TRAT;

TITLE2 'MODELO EFEITOS PRINCIPAIS C/INTERAÇÃO-RESPOSTA LOGITO';

RESPONSE 1 -1 LOG;

MODEL ESC = PESO TRAT PESO*TRAT/ PRED;

RUN;

TITLE2 'MODELO IDENTIDADE - RESPOSTA LOGITO';

MODEL ESC=(1 0 0 0 0 0 0 0 0 0 0 0 0,

0 1 0 0 0 0 0 0 0 0 0 0 0,

0 0 1 0 0 0 0 0 0 0 0 0 0,

0 0 0 1 0 0 0 0 0 0 0 0 0,

0 0 0 0 1 0 0 0 0 0 0 0 0,

0 0 0 0 0 1 0 0 0 0 0 0 0,

0 0 0 0 0 0 1 0 0 0 0 0 0,

0 0 0 0 0 0 0 1 0 0 0 0 0,

0 0 0 0 0 0 0 0 1 0 0 0 0,

0 0 0 0 0 0 0 0 0 1 0 0 0,

0 0 0 0 0 0 0 0 0 0 1 0 0,

0 0 0 0 0 0 0 0 0 0 0 1 0,

0 0 0 0 0 0 0 0 0 0 0 0 1)

(1='C1 SR',

2='C1 CR',

3='C2 SR',

4='C2 CR',

5='C3 SR',

6='C3 CR',

7='C4 SR',

8='C4 CR',

9='C5 SR',

10='C5 CR',

11='C6 SR',

12='C6 CR')/NODESIGN NOPROFILE PRED;

RUN;

CONTRAST 'NAO INTERACAO PESO X TRATAM' ALL_PARMS 1 -1 -1 1 0 0
0 0 0 0 0 0,
ALL_PARMS 1 -1 0 0 -1 1
0 0 0 0 0 0,
ALL_PARMS 1 -1 0 0 0 0
-1 1 0 0 0 0,
ALL_PARMS 1 -1 0 0 0 0
0 0 -1 1 0 0,

```
ALL_PARMS 1 -1 0 0 0 0
           0 0 0 0 -1 1;
```

```
RUN;
```

```
TITLE2 'MODELO DE EFEITOS PRINCIPAIS';
```

```
MODEL ESC=PESO TRAT /PRED NOPROFILE;
```

```
RUN;
```

```
MODEL ESC=(1 1 0 0 1,
            1 1 0 0 -1,
            1 0 0 0 1,
            1 0 0 0 -1,
            1 0 1 0 1,
            1 0 1 0 -1,
            1 0 0 1 1,
            1 0 0 1 -1,
            1 0 0 0 1,
            1 0 0 0 -1,
            1 -1 -1 -1 1,
            1 -1 -1 -1 -1)/NODESIGN NOPROFILE PRED;
```

```
RUN;
```

TITLE 'ESTRUTURA II - RESPOSTA ORDINAL POLITOMICA';

DATA ABCX21;

INPUT PESO \$ TRAT \$ ESC \$ CONT @@;

CARDS;

C1 SR A 4 C3 SR A 3 C5 SR A 1
C1 SR M 6 C3 SR M 5 C5 SR M 3
C1 SR B 30 C3 SR B 32 C5 SR B 36

C1 CR A 8 C3 CR A 5 C5 CR A 4
C1 CR M 10 C3 CR M 6 C5 CR M 5
C1 CR B 22 C3 CR B 29 C5 CR B 31

C2 SR A 7 C4 SR A 3 C6 SR A 2
C2 SR M 8 C4 SR M 3 C6 SR M 3
C2 SR B 25 C4 SR B 34 C6 SR B 35

C2 CR A 9 C4 CR A 4 C6 CR A 4
C2 CR M 10 C4 CR M 5 C6 CR M 3
C2 CR B 21 C4 CR B 31 C6 CR B 33

;

PROC CATMOD ORDER=DATA;

WEIGHT CONT;

POPULATION PESO TRAT;

TITLE2 'MODELO EFEITOS PRINCIPAIS COM INTERAÇÃO-RESPOSTA ESCORE MEDIO';

RESPONSE 0.75 0.250 0;

MODEL ESC = PESO TRAT PESO*TRAT/COV PRED;

RUN;

TITLE2 'MODELO IDENTIDADE - RESPOSTA ESCORE MEDIO';

MODEL ESC=(1 0 0 0 0 0 0 0 0 0 0 0 0,
0 1 0 0 0 0 0 0 0 0 0 0 0,
0 0 1 0 0 0 0 0 0 0 0 0 0,
0 0 0 1 0 0 0 0 0 0 0 0 0,
0 0 0 0 1 0 0 0 0 0 0 0 0,
0 0 0 0 0 1 0 0 0 0 0 0 0,
0 0 0 0 0 0 1 0 0 0 0 0 0,
0 0 0 0 0 0 0 1 0 0 0 0 0,
0 0 0 0 0 0 0 0 1 0 0 0 0,
0 0 0 0 0 0 0 0 0 1 0 0 0,
0 0 0 0 0 0 0 0 0 0 1 0 0,
0 0 0 0 0 0 0 0 0 0 0 1 0,
0 0 0 0 0 0 0 0 0 0 0 0 1)

(1='C1 SR',
2='C1 CR',
3='C2 SR',
4='C2 CR',
5='C3 SR',
6='C3 CR',
7='C4 SR',
8='C4 CR',
9='C5 SR',
10='C5 CR',
11='C6 SR',
12='C6 CR')/NODESIGN COV PRED;

RUN;

```

CONTRAST 'NAO INTERACAO PESO X TRATAM' ALL_PARMS 1 -1 -1 1 0 0
          0 0 0 0 0 0,
          ALL_PARMS 1 -1 0 0 -1 1
          0 0 0 0 0 0,
          ALL_PARMS 1 -1 0 0 0 0
          -1 1 0 0 0 0,
          ALL_PARMS 1 -1 0 0 0 0
          0 0 -1 1 0 0,
          ALL_PARMS 1 -1 0 0 0 0
          0 0 0 0 -1 1;

```

```

RUN;
TITLE2 'MODELO DE EFEITOS PRINCIPAIS - RESPOSTA ESCORE MEDIO';
MODEL ESC=PESO TRAT /COV PRED;

```

```

RUN;
MODEL ESC=(1 1 0 0 1,
           1 1 0 0 -1,
           1 0 0 0 1,
           1 0 0 0 -1,
           1 0 1 0 1,
           1 0 1 0 -1,
           1 0 0 1 1,
           1 0 0 1 -1,
           1 0 0 0 1,
           1 0 0 0 -1,
           1 -1 -1 -1 1,
           1 -1 -1 -1 -1)/NODESIGN NOPROFILE PRED;

```

```

RUN;

```



```

LIBNAME AHM 'B:\';
PROC IML WORKSIZE=1000;
p={0.80, 0.89, 0.83, 0.84,
   0.67, 0.80, 0.72, 0.70,
   0.64, 0.77, 0.67, 0.69};

p1={0.80, 0.80, 0.80};
p2={0.67, 0.67, 0.67};
p3={0.64, 0.64, 0.64};

p124={0.89, 0.83, 0.84};
p224={0.80, 0.72, 0.70};
p324={0.77, 0.67, 0.69};

r={1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1};

r1={1, 1, 1};

q=r-p; Dp=DIAG(p); Dq=DIAG(q);

/*CALCULO DA MATRIZ DE COVARIANCIAS DAS ESTIMATIVAS*/;

Vp=(Dp*Dq);
dVp=VECDIAG(Vp);

/*CALCULO DA MATRIZ DE COVARIANCIAS DE F(p)*/;

A1={1 -1 0 0,
     1 0 -1 0,
     1 0 0 -1};

A=BLOCK(A1,A1,A1);
d=A*P;

A11= {1 1 1 0 0 0 0 0 0};
A12= {0 0 0 1 1 1 0 0 0};
A13= {0 0 0 0 0 0 1 1 1};

d1= A11*d;
d2= A12*d;
d3= A13*d;

Fp=(EXP(-5*(INV(DIAG(P1//P2//P3))*d)##2));

F1=A11*Fp;
F2=A12*Fp;
F3=A13*Fp;

PRINT Fp;

M1=-10*(INV(DIAG(p1))*D1)*(R1+(INV(DIAG(P1##2)))*P124)*F1;
M2=-10*(INV(DIAG(p2))*D2)*(R1+(INV(DIAG(P2##2)))*P224)*F2;

```

```
M3=-10*(INV(DIAG(p3))*D3)*(R1+(INV(DIAG(P3##2)))*P324)*F3;
```

```
C1=-10*(INV(DIAG(p1))*D1)*(R1-(INV(DIAG(P1)))*P124)*F1;
```

```
C2=-10*(INV(DIAG(p2))*D2)*(R1-(INV(DIAG(P2)))*P224)*F2;
```

```
C3=-10*(INV(DIAG(p3))*D3)*(R1-(INV(DIAG(P3)))*P324)*F3;
```

```
D1=M1|| (DIAG(c1));
```

```
D2=M2|| (DIAG(c2));
```

```
D3=M3|| (DIAG(c3));
```

```
PRINT d1;
```

```
H=BLOCK(D1,D2,D3);
```

```
VFp=(1/200)*(H*Vp*H`);
```

```
FpVFP=Fp`//VFp;
```

```
CREATE COVSORG FROM FpVFP;
```

```
APPEND FROM FpVFP;
```

```
QUIT;
```

```
TITLE 'ANALISE DA VARIACAO DA EFICIENCIA DOS METODOS ';
```

```
DATA COVSORG(TYPE=EST);
```

```
SET COVSORG;
```

```
IF _N_=1 THEN DO;
```

```
  _TYPE_='PARMS';
```

```
  _NAME_='.';
```

```
  B1=COL1;B2=COL2;B3=COL3;B4=COL4;B5=COL5;B6=COL6;
```

```
  B7=COL7;B8=COL8;B9=COL9;
```

```
OUTPUT;
```

```
END;
```

```
ELSE DO;
```

```
  _TYPE_='COV';
```

```
IF _N_=2 THEN _NAME_='B1';
```

```
IF _N_=3 THEN _NAME_='B2';
```

```
IF _N_=4 THEN _NAME_='B3';
```

```
IF _N_=5 THEN _NAME_='B4';
```

```
IF _N_=6 THEN _NAME_='B5';
```

```
IF _N_=7 THEN _NAME_='B6';
```

```
IF _N_=8 THEN _NAME_='B7';
```

```
IF _N_=9 THEN _NAME_='B8';
```

```
IF _N_=10 THEN _NAME_='B9';
```

```
  B1=COL1;B2=COL2;B3=COL3;B4=COL4;B5=COL5;B6=COL6;
```

```
  B7=COL7;B8=COL8;B9=COL9;
```

```
OUTPUT;
```

```
END;
```

```
DATA COVSORG;
```

```
SET COVSORG;
```

```
KEEP B1-B9 _TYPE_ _NAME_;
```

```
PROC PRINT;
```

```
PROC CATMOD DATA=COVSORG;
```

```
RESPONSE READ B1-B9;
```

```
FACTORS CULT $ 3, TRAT $ 3 / _RESPONSE_= CULT TRAT(CULT)
```

```
PROFILE =(A 'GERMINACAO',
```

```
  A 'TZVIGOR' ,
```

A 'IMERSAO',
B 'GERMINACAO' ,
B 'TZVIGOR' ,
B 'IMERSAO',
C 'GERMINACAO' ,
C 'TZVIGOR',
C 'IMERSAO');

MODEL _F=_RESPONSE_/PRED COV COVB TITLE='MODELO DE EFEITOS HIERARQUICOS
RUN;

CONTRAST 'NAO INTERACAO CULT VS TRAT' ALL_PARMS 0 0 1 0 -1 0 0 0,
ALL_PARMS 0 0 1 0 0 0 -1 0,
ALL_PARMS 0 0 0 1 0 -1 0 0,
ALL_PARMS 0 0 0 1 0 0 0 -1;

CONTRAST 'TRAT DENTRO DE CULT 1' ALL_PARMS 0 0 1 0 0 0 0 0,
ALL_PARMS 0 0 0 1 0 0 0 0;

CONTRAST 'TRAT DENTRO DE CULT 2' ALL_PARMS 0 0 0 0 1 0 0 0,
ALL_PARMS 0 0 0 0 0 1 0 0;

CONTRAST 'TRAT DENTRO DE CULT 3' ALL_PARMS 0 0 0 0 0 0 1 0,
ALL_PARMS 0 0 0 0 0 0 0 1;

RUN;

```

LIBNAME AHM 'B:\';
PROC IML WORKSIZE=1000;
p={0.80, 0.89, 0.83, 0.84,
    0.67, 0.80, 0.72, 0.70,
    0.64, 0.77, 0.67, 0.69};

p1={0.80, 0.80, 0.80};
p2={0.67, 0.67, 0.67};
p3={0.64, 0.64, 0.64};

p124={0.89, 0.83, 0.84};
p224={0.80, 0.72, 0.70};
p324={0.77, 0.67, 0.69};

r={1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1};

r1={1, 1, 1};

q=r-p; Dp=DIAG(p); Dq=DIAG(q);

/*CALCULO DA MATRIZ DE COVARIANCIAS DAS ESTIMATIVAS*/;

Vp=(Dp*Dq);
dVp=VECDIAG(Vp);

/*CALCULO DA MATRIZ DE COVARIANCIAS DE F(p)*/;

A1={1 -1 0 0,
    1 0 -1 0,
    1 0 0 -1};

A=BLOCK(A1,A1,A1);
C=A*P;

A11= {1 1 1 0 0 0 0 0 0};
A12= {0 0 0 1 1 1 0 0 0};
A13= {0 0 0 0 0 0 1 1 1};

d1= A11*d;
d2= A12*d;
d3= a13*d;

Fp=(EXP(-5*(INV(DIAG(P1//P2//P3))*d)))#2);

F1=A11*Fp;
F2=A12*Fp;
F3=A13*Fp;

M1=-10*(INV(DIAG(p1))*D1)*(R1+(INV(DIAG(P1##2)))*P124)*F1;
M2=-10*(INV(DIAG(p2))*D2)*(R1+(INV(DIAG(P2##2)))*P224)*F2;
M3=-10*(INV(DIAG(p3))*D3)*(R1+(INV(DIAG(P3##2)))*P324)*F3;

```

```
C1=-10*(INV(DIAG(p1))*D1)*(R1-(INV(DIAG(P1)))*P124)*F1;
C2=-10*(INV(DIAG(p2))*D2)*(R1-(INV(DIAG(P2)))*P224)*F2;
C3=-10*(INV(DIAG(p3))*D3)*(R1-(INV(DIAG(P3)))*P324)*F3;
```

```
D1=M1|| (DIAG(c1));
D2=M2|| (DIAG(c2));
D3=M3|| (DIAG(c3));
```

```
H=BLOCK(D1,D2,D3);
```

```
VFp=(1/200)*(H*Vp*H`);
```

```
FpVFP=Fp`//VFp;
```

```
CREATE COVSORG FROM FpVFp;
```

```
APPEND FROM FpVFp;
```

```
QUIT;
```

```
TITLE 'ANALISE DA VARIACAO DA EFICIENCIA DOS METODOS ';
```

```
DATA COVSORG(TYPE=EST);
```

```
SET COVSORG;
```

```
IF _N_=1 THEN DO;
```

```
  _TYPE_='PARMS';
```

```
  _NAME_='.';
```

```
  B1=COL1;B2=COL2;B3=COL3;B4=COL4;B5=COL5;B6=COL6;
```

```
  B7=COL7;B8=COL8;B9=COL9;
```

```
OUTPUT;
```

```
END;
```

```
ELSE DO;
```

```
  _TYPE_='COV';
```

```
IF _N_=2 THEN _NAME_='B1';
```

```
IF _N_=3 THEN _NAME_='B2';
```

```
IF _N_=4 THEN _NAME_='B3';
```

```
IF _N_=5 THEN _NAME_='B4';
```

```
IF _N_=6 THEN _NAME_='B5';
```

```
IF _N_=7 THEN _NAME_='B6';
```

```
IF _N_=8 THEN _NAME_='B7';
```

```
IF _N_=9 THEN _NAME_='B8';
```

```
IF _N_=10 THEN _NAME_='B9';
```

```
  B1=COL1;B2=COL2;B3=COL3;B4=COL4;B5=COL5;B6=COL6;
```

```
  B7=COL7;B8=COL8;B9=COL9;
```

```
OUTPUT;
```

```
END;
```

```
DATA COVSORG;
```

```
SET COVSORG;
```

```
KEEP B1-B9 _TYPE_ _NAME_ ;
```

```
PROC PRINT;
```

```
PROC CATMOD DATA=COVSORG;
```

```
RESPONSE READ B1-B9;
```

```
FACTORS CULT $ 3, TRAT $ 3 / _RESPONSE_= CULT TRAT
```

```
PROFILE =(A 'GERMINACAO',
```

```
  A 'TZVIGOR' ,
```

```
  A 'IMERSAO',
```

```
  B 'GERMINACAO' ,
```

```
B 'TZVIGOR' ,  
B 'IMERSAO',  
C 'GERMINACAO' ,  
C 'TZVIGOR',  
C 'IMERSAO');
```

```
MODEL F=_RESPONSE_/PRED COV COVB TITLE='MODELO DE EFEITOS PRINCIPAIS.'
```

```
RUN;
```

```
CONTRAST 'AUSENCIA EF. METODOS' ALL_PARMS 0 0 0 1 0,  
ALL_PARMS 0 0 0 0 1;
```

```
RUN;
```

```

LIBNAME AHM 'B:\';
PROC IML WORKSIZE=1000;
Y={119,84,83};
F=(1/14)*Y;
DF=DIAG(F);
VF=((1/14)**2)*DIAG(F);
FVF=F`//VF;

CREATE COVRAIZ FROM FVF;
APPEND FROM FVF;
QUIT;
TITLE 'EFEITO DO SUBSTRATO NO ENRAIZAMENTO DE ALPORQUES DE URUCUM';
DATA COVRAIZ(TYPE=EST);
SET COVRAIZ;
  IF _N_=1 THEN DO;
    _TYPE_='PARMS';
    _NAME_='.';
    B1=COL1;B2=COL2;B3=COL3;
  OUTPUT;
  END;
  ELSE DO;
    _TYPE_='COV';
    IF _N_=2 THEN _NAME_='B1';
    IF _N_=3 THEN _NAME_='B2';
    IF _N_=4 THEN _NAME_='B3';

    B1=COL1;B2=COL2;B3=COL3;
  OUTPUT;
  END;
DATA COVRAIZ;
SET COVRAIZ;
KEEP B1-B3 _TYPE_ _NAME_ ;
PROC PRINT;
PROC CATMOD DATA=COVRAIZ;
TITLE2 'MODELO DE EFEITOS PRINCIPAIS';
RESPONSE READ B1-B3;
FACTORS TRAT $ 3 / _RESPONSE_ = TRAT
PROFILE =('AREIA',
          'BARRO',
          '1/2A+1/2B');

MODEL _F_=(1 0 0,
          0 1 0,
          0 0 1)(1='MEDIA DE A',
                2='MEDIA DE B',
                3='MEDIA DE C')/PRED NOINT COVB
TITLE2='MODELO IDENTIDADE SATURADO';
RUN;
CONTRAST 'MEDIA ARITMETICA' ALL_PARMS -1 -1 2;
CONTRAST 'SUBS 1/2A POR 1/2B' ALL_PARMS 1 0 -1;
CONTRAST 'SUBS 1/2B POR 1/2A' ALL_PARMS 0 1 -1;
RUN;
MODEL _F_=(1 0,
          0 1,

```

0 1) (1='MEDIA DE A',
2='MEDIA DE B OU C') / COVB PRED

TITLE='MODELO REDUZIDO';

RUN;

CONTRAST 'AUSENCIA EF. TRAT' ALL_PARMS 1 -1;

RUN;

Refêrencias Bibliográficas

- [1] Arnold, S. F.. *The theory of linear models and multivariate analysis*. New York, John Wiley, 1981.
- [2] Bickel, P. J. & Doksum, K. A.. *Mathematical statistics: basic ideas and select topics*. San Francisco, Ca., Holden Day Inc, 1977.
- [3] Bishop, Y. M. M.; Fienberg, S. E.; Holland, P.W.. *Discrete multivariate analysis: theory and practice*, Cambridge, MA: The MIT Press, 1975.
- [4] Botrel, N. & Carvalho, V. D.. Efeito do peso do fruto no escurecimento interno e qualidade do abacaxi Smooth Cayenne. *Pesq. agropec. bras.*, Brasília. 28(9): 1055-64. setembro, 1993.
- [5] Fienberg, S. E.. *The analysis of cross classified categorical data*. 2nd ed.. Cambridge, MA: The MIT Press, 1980.
- [6] Forthofer, R.N. & Koch, G.G.. An analysis of compounded functions of categorical data, *Biometrics*, 29, 143-157, 1973.
- [7] Forthofer, R. N. & Lehnen, R. G.. *Public program analysis: a new categorical data approach*. Belmont, CA, Wadsworth, 1981.
- [8] Freeman, D. H.. *Applied categorical data analysis*. New York. Marcel Dekker Inc., 1987.
- [9] Grizzle, J.E.; Starmer, C. F.; Koch, G. G.. Analysis of categorical data by linear models, *Biometrics*, 25, 489-504.
- [10] Haberman, S. J.. *Analysis of qualitative data*. New York: Academic Press Inc..1978, 1979.
- [11] Halmos, P. R.. *Finite dimensional vector spaces*. 2nd ed, New York, Van Nostrand, 1958.

- [12] Hoffman, K. & Kunze, R.. *Álgebra linear*. São Paulo, EDUSP, 1971.
- [13] James, B. R.. *Probabilidade: um curso em nível intermediário*. Rio de Janeiro, IMPA, 1981.
- [14] Joshi, D. D.. *Linear estimation and design of experiments*. New York, Wiley, 1971.
- [15] Koch, G. C.; Imrey, P.B.; Singer, J. M.; Atkinson, S. S.; Stokes, M.E.. *Lectures notes for analysis of categorical data*. 1984.
- [16] Leite, J. G. & Singer, J. da M.. *Métodos assintóticos em estatística: fundamentos e aplicações*. São Paulo, IX SINAPE, 1990.
- [17] Montgomery, D. C. & Peck, E.. *Introduction to linear regression analysis*. New York, Wiley, 1982.
- [18] Mota, R. M. S.. *Caracterização dos melhores estimadores não tendenciosos no modelo linear geral*. USP, São Paulo, 1982. (Tese de Mestrado).
- [19] Neyman, J.. Contribution to theory of the χ^2 - test. *Proc. 1st Berkeley Symposium on Math. Statist. and Prob.*, University of California Press, Berkeley, pp. 239-73. 1949.
- [20] Petrini, J. A.. Testes de vigor para prever emergência em sorgo. *Pesq. agropec. bras.*, Brasília, 23(6): 593-98, setembro, 1988.
- [21] Plackett, R. L.. *The analysis of categorical data*. High Wycombe, Buckinghamshire, England: Charles Griffin & Company Ltda, 1974.
- [22] SAS Institute Inc. *SAS/IMLTM Users Guide, Version 5 edition*. Cary, NC: SAS Institute Inc, 1985, 300pp.
- [23] SAS Institute Inc. *SAS/STATTM Users Guide, Version 6.3 edition*. Cary, NC: SAS Institute Inc, 1988, 1028pp.
- [24] Searle, R. S.. *Introduction to linear model*. New York, Wiley, 1971.
- [25] Silva, K. M. B. e ; Almeida, F. C. G. ; Almeida, F. A. G. ; Silva, P. S. L. e ; Albuquerque, J. J. L.. Efeito do substrato no enraizamento de alporques de urucuzeiro. *Pesq. agropec. bras.*, Brasília, 28(1): 101-106, janeiro, 1993.

- [26] Singer, J. da M. & Leite, J. G.. Métodos assintóticos em estatística: fundamentos e aplicações. São Paulo, IX SINAPE, 1990.
- [27] Wald, A.. Tests of statistical hypothesis concerning several parameters when the number of observations is large. *Transaction of the American Mathematical Society*. 54: 426-82, 1943.