



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E MATEMÁTICA APLICADA
CURSO DE ESTATÍSTICA

LAURA VICUÑA TORRES DE PAULA

**AVALIANDO A ACEITAÇÃO DA IMPRESSÃO GLOBAL DO SUCO DE
CAJU VIA MODELOS DE CHANCES PROPORCIONAIS MISTOS**

FORTALEZA
2013

LAURA VICUÑA TORRES DE PAULA

AVALIANDO A ACEITAÇÃO DA IMPRESSÃO GLOBAL DO SUCO DE
CAJU VIA MODELOS DE CHANCES PROPORCIONAIS MISTOS

Monografia apresentada ao curso de Estatística do Departamento de Estatística e Matemática Aplicada da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Bacharel em Estatística.

Orientadora: Profa. Dra. Ana Maria Souza de Araujo

FORTALEZA

2013

LAURA VICUÑA TORRES DE PAULA

AVALIANDO A ACEITAÇÃO DA IMPRESSÃO GLOBAL DO SUCO DE
CAJU VIA MODELOS DE CHANCES PROPORCIONAIS MISTOS

Monografia apresentada ao curso de Estatística do Departamento de Estatística e Matemática Aplicada da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Bacharel em Estatística.

Aprovada em ___/___/_____.

BANCA EXAMINADORA

Profa. Dra. Ana Maria Souza de Araujo (Orientadora)
Universidade Federal do Ceará (UFC)

Profa. Dra. Sílvia Maria de Freitas
Universidade Federal do Ceará (UFC)

Prof. Dr. Ronald Targino Nojosa
Universidade Federal do Ceará (UFC)

Dedico esse trabalho

À Deus acima de tudo

À minha mãe Engrácia, por ser a melhor mãe do mundo

À meus irmãos, Emanuel, Emanuela, Vivianny e Luiz Antônio.

“A educação é a arma mais poderosa que você pode usar para mudar o mundo.”

Nelson Mandela

“A educação tem raízes amargas, mas os seus frutos são doces.”

Aristóteles

AGRADECIMENTOS

Agradeço primeiramente a Deus, por sempre iluminar meu caminho com saúde, sabedoria, determinação e por guiar pessoas especiais para minha vida, que sempre me apoiaram nas minhas escolhas e estiveram comigo nos momentos bons e ruins.

À minha mãe Engrácia, por ter me dado a oportunidade de estar aqui hoje. Pelo seu esforço para que nunca me faltasse nada, sua força, carinho e dedicação. Por seu apoio incondicional. Agradeço pelos seus sorrisos e lágrimas de alegria e satisfação que estarão sempre gravados em meu coração. Agradeço pelo simples e completo “Eu te amo” todas as noites ao falar comigo pelo telefone. Mãe, você é a melhor mãe do mundooooo!!! Eu te amo e obrigada por tudo.

À minha tia Lúcia, por ter sempre me ajudado para que eu pudesse estar aqui hoje, por seus ensinamentos, carinho e conselhos. Sou muito grata por tudo.

À minha orientadora, professora Ana Maria, por sua orientação acadêmica e paciência na elaboração desse trabalho, sou muito grata pelo apoio, confiança e compreensão, principalmente nos momentos difíceis. Agradeço muito a Deus por ter colocado a senhora no meu caminho. Obrigada pela oportunidade de poder ser sua orientanda mais uma vez e pela sua calma e tranquilidade nesse período de 1 ano e 6 meses para a elaboração desse trabalho.

Agradeço aos meus irmãos Emanuel, Emanuela, Vivianny e Luiz Antônio pelo carinho, disposição e toda força dada. Por ser meu amigo ou amiga e tudo mais que uma pessoa pode ser à outra.

Aos meus sobrinhos lindos, Ana Luzia, José Guilherme e Maria Tereza, por serem crianças adoráveis e carinhosas. Por me terem em seus corações e por me permitirem fazer parte de suas vidas.

Agradeço a todos os professores do Departamento de Estatística e Matemática Aplicada, principalmente aqueles que puderam ser meus professores em alguma disciplina acadêmica do curso, pois pude aprender muito com vocês, não só no estudo mais como pessoa também.

Às funcionárias Luiza, Márgeri e Dona Margarida que sempre tiveram muita paciência e simpatia por mim, e por sempre me ajudar, quando estava ao seu alcance. Meu muito obrigada.

Ao grupo “Família” da empresa júnior de Estatística GAUSS que sempre me apoio nas minhas escolhas. Por todos os ensinamentos adquiridos nesses período de 2 anos, onde estes me acrescentou muito como estudante e como pessoa. Muito obrigada por tudo.

Às grandes amigas Janaína MLR, Kelly MLR e Raquel MLR que sempre estiveram comigo nesses 4 anos e 6 meses, tanto nos momentos felizes como tristes. Agradeço por sua ajuda, por sempre estender a mão sempre que possível, pela paciência e apoio em todos os momentos. Eu nunca esquecerei das longas tarde de risadas, “Porque se for pra ir agora é melhor ir depois”, “Por que ou você ganha ou você perde” e “Desespero da Kelly ao saber de uma prova marcada de uma disciplina sendo que não tinha prova marcada nenhuma, era apenas um trote da Janaína” (rsrsrsr). Obrigada por suas amizades incondicionais.

RESUMO

O modelo de chances proporcionais é um método que pode ser utilizado para descrever a relação entre uma variável resposta categórica ordinal e uma ou mais covariáveis. Sendo este uma extensão dos modelos lineares generalizados, ele permite que se obtenham probabilidades acumuladas para cada uma das categorias de resposta e, por consequência, as probabilidades marginais. Com isso, o presente trabalho tem como finalidade ajustar um modelo de chances proporcionais para verificar a aceitação da impressão global do suco de caju. As possíveis respostas para essa variável são categorias ordinais para cinco diferentes marcas de suco de caju. Um modelo de chances proporcionais foi ajustado para esse experimento e, a fim de tratar uma possível dependência existente entre as observações, foi adicionado no modelo um efeito aleatório.

O *software* R foi utilizado para estimar os modelos, elaborar os gráficos cruzados e os de probabilidades acumuladas da variável **impressão global**, no qual, foi usado o pacote **ordinal**, com a função `clm` e `clmm2`.

Palavras-chave: Efeito Aleatório, Modelos de Chances Proporcionais, Modelos Lineares Generalizados.

ABSTRACT

The proportional odds model is a method that can be used to describe the relationship between ordinal categorical variables. Since this is an extension of generalized linear models, it allows to obtain cumulative probabilities for each of the response categories and, therefore, the marginal probabilities. Thus, this study aims to fit a proportional odds model to verify acceptance of the overall impression of the cashew juice. The possible answers to this variable are ordinal categories for five different brands of cashew juice. A model was adjusted for this experiment, and in order to model the possible dependence between the observations a random effect was added to the proportional odds model.

Keywords: Generalized Linear Models, Random Effect, Proportional odds models.

LISTA DE FIGURAS

1	Modelo de chances proporcionais para a escala do log <i>odds</i>	32
2	Comportamento da função de probabilidade do modelo de chances proporcionais	33
3	Cruzamento das variáveis marca, consumo por suco de caju e impressão global.	52
4	Cruzamento das variáveis marca, consumo por outros sucos de frutas e impressão global.	52
5	Cruzamento das variáveis marca, gosto por suco de caju e impressão global.	53
6	Cruzamento das variáveis marca, gosto por outros sucos de frutas e impressão global	53
7	Estimativas das probabilidades referentes ao modelo ajustado sem o efeito aleatório. 55	
8	Verossimilhança perfilada de σ_u do modelo (4.1) para os dados de suco de caju.	56
9	Estimativas das probabilidades referentes ao modelo ajustado sem o efeito aleatório. 57	

LISTA DE TABELAS

1	Funções de ligação canônicas	19
2	Mulheres com câncer de mama	29
3	Estimativas para os parâmetros do modelo de chances proporcionais para impressão global do suco de caju.	54
4	Estimativas para os parâmetros do modelo de chances proporcionais para impressão global do suco de caju.	57
5	Cruzamento das variáveis impressão global, marca e consumo de suco de caju.	63
6	Cruzamento das variáveis impressão global, marca e consumo por outros sucos de frutas.	64
7	Cruzamento das variáveis impressão global, marca e gosto de suco de caju.	65
8	Cruzamento das variáveis impressão global, marca e gosto por outros sucos de frutas.	66

SUMÁRIO

1	INTRODUÇÃO	14
2	MODELO LINEAR GENERALIZADO	16
2.1	Definição	17
2.2	Estimação	19
2.3	Comparação de modelos	20
2.3.1	Critérios AIC e BIC	21
2.3.2	<i>Deviance</i>	21
2.4	Testes de Hipótese	22
2.4.1	Razão de verossimilhança	23
2.4.2	Teste de Wald	23
2.4.3	Teste Escore	24
3	MODELOS PARA VARIÁVEIS QUALITATIVAS	25
3.1	Modelo logístico	26
3.1.1	Definição	26
3.1.2	Estimação	27
3.1.3	<i>Odds Ratio</i>	28
3.1.3.1	Intervalo de confiança para o <i>odds ratio</i>	30
3.1.3.2	Teste de hipótese para <i>odds ratio</i>	30
3.2	Modelo de chances proporcionais (MCP)	30
3.2.1	Definição	31
3.2.2	Estimação	33

4	MODELOS MISTOS	35
4.1	Modelo linear misto clássico	35
4.1.1	Estimação	36
4.2	Modelo linear generalizado misto (MLGM)	38
4.3	MCP com efeito aleatório	39
4.3.1	Estimação	40
4.3.1.1	Função de verossimilhança para o efeito fixo	40
4.3.1.2	Função de verossimilhança para o efeito aleatório	42
4.3.2	Testes de Hipótese	42
4.3.2.1	Teste da razão de verossimilhança	42
4.3.2.2	Teste de Wald	43
4.3.2.3	Teste de Escore	43
4.3.3	Verossimilhança perfilada	44
4.3.3.1	Função	44
5	ANÁLISE DE DADOS DE SUCO DE CAJU	49
5.1	Pacote Ordinal	50
5.2	Análise descritiva	51
5.3	Modelo de efeitos fixos	51
5.4	Modelo de efeitos aleatórios	55
6	CONSIDERAÇÕES FINAIS	58
	REFERÊNCIAS	59
	APÊNDICE	62
	ANEXO	82

1 INTRODUÇÃO

O modelo de regressão de chances proporcionais (“proportional odds”) é uma metodologia que pode ser utilizada para descrever a relação entre variáveis categorizadas ordinais e uma ou mais covariáveis (McCULLAGH, 1980), permitindo que se obtenham probabilidades acumuladas para cada uma das categorias de resposta e, por consequência, as probabilidades marginais.

Há estudos em que as variáveis ordinais são avaliadas em mais de uma ocasião, ocorrendo mais de uma medida na mesma unidade experimental, caracterizando um experimento com medidas repetidas que são aqueles em que cada unidade experimental é observada em pelo menos duas ocasiões. Por esse motivo espera-se uma dependência entre as observações referentes à mesma unidade experimental (NOBRE, 2004). Pode-se citar, como exemplo, experimentos em Tecnologia de Alimentos, nos quais avaliadores provam vários produtos e os avaliam, sendo a avaliação realizada por meio da escala de Likert. Define-se como escala de Likert como sendo não comparativa e do tipo itemizada, que pode ser utilizada para avaliar produtos/serviços, onde os entrevistados assinalam um único item de acordo com seu grau de satisfação (CUNHA, 2007). Stram, Wei e Ware (1988) propuseram o ajuste de modelos marginais com dados dessa natureza para os casos em que se tem mais de uma ocasião de medida (tempo, dosagem, espaçamento). Uma alternativa para tratar a possível dependência existente entre as observações é a utilização de modelos mistos (COSTA, 2012). Estes modelos são caracterizados pela introdução de um efeito aleatório associado a cada observação, seguindo uma distribuição pré-especificada, permitindo acrescentar uma estrutura de covariância entre as medidas de uma mesma unidade experimental.

O presente trabalho tem por objetivo utilizar o modelo de chances proporcionais com efeitos aleatórios para verificar uma possível relação entre a variável impressão global e algumas covariáveis em um experimento, no qual cinco marcas de suco de caju são avaliadas segundo uma escala ordinal de resposta (Likert). Foram ajustados dois modelos para variável impressão global. O primeiro modelo foi o de chances proporcionais com efeitos fixos e o segundo, o modelo com efeitos aleatórios, fazendo-se uma comparação

entre os resultados dos dois modelos.

O trabalho está organizado da seguinte forma: no Capítulo 2 serão abordados resultados básicos de modelos lineares generalizados, envolvendo estimação, comparação de modelos e testes de hipóteses; no Capítulo 3 serão abordados métodos de modelos para variáveis qualitativas, envolvendo o modelo logístico, o modelo de chances proporcionais e suas respectivas estimativas; no Capítulo 4 serão abordados os modelos mistos, envolvendo o modelo linear misto clássico, o modelo linear generalizado misto, o modelo de chances proporcionais com efeito aleatório, suas respectivas estimativas, testes de hipóteses e verossimilhança perfilada. A análise descritiva dos dados e o ajuste dos modelos foi feito utilizando-se o *software* R 2.13.1. Os dados são resultados de uma pesquisa referentes a uma tese de doutorado do programa pertencente à Rede Nordeste de Biotecnologia (RENORBIO) (LEMOS, 2010). Por fim, serão apresentadas ainda as considerações finais sobre esse estudo.

2 *MODELO LINEAR GENERALIZADO*

Muitos estudos estatísticos referem-se ao problema de verificar a influência de uma ou mais variáveis explicativas sobre uma variável resposta. Para verificar esta relação é esperado que a variável resposta e as variáveis explicativas (também chamadas de co-variáveis) sejam pelo menos correlacionadas, para que as variáveis explicativas contenham informações sobre a variável resposta. Esse tipo de problema pode ser resolvido pelos modelos de regressão.

Segundo Hoffmann (2006), o modelo de regressão linear clássico pode ser representado por

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.1)$$

sendo que $\mathbf{y}_{n \times p}$ é o vetor de observações da variável resposta, $\mathbf{X}_{n \times p}$ é a matriz de observações das variáveis explicativas (matriz do modelo), $\boldsymbol{\beta}_{p \times 1}$ o vetor de parâmetros do modelo e por último o vetor de erros aleatórios representado por $\boldsymbol{\varepsilon}_{n \times 1}$. Algumas pressuposições são feitas sobre o modelo de regressão linear clássico. Supõe-se que:

1. A relação entre \mathbf{X} e \mathbf{y} seja linear;
2. O vetor de erros seja 0, isto é, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$;
3. A variância dos erros seja sempre σ^2 , ou seja, os erros sejam homocedásticos;
4. Os erros devem ter distribuição Normal, ou seja, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}_n, \sigma^2 \mathbf{I}_n)$, em que \mathbf{I}_n representa a matriz identidade de ordem n ;
5. O erro de uma observação seja não correlacionado com o erro de outra observação, isto é, $E(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i)E(\varepsilon_j)$.

Em alguns casos, a suposição de normalidade não é atendida pela variável resposta. Assim, foi surgindo na literatura estatística outros modelos, como por exemplo: modelo

complemento log-log, modelo probit, modelo logit, modelo log-linear, dentre outros. Todos esses exemplos são originários de uma estrutura não linear em um conjunto linear de parâmetros, mas que são linearizados. Além disso, a variável resposta não segue distribuição normal e sim segue a distribuição específica de cada modelo. Estes casos podem ser tratados por meio da teoria de Modelos Lineares Generalizados (MLG).

O MLG tem como principal característica o fato da variável resposta pertencer à família exponencial na forma canônica, abrangendo assim os modelos citados anteriormente. A utilização desse modelo proposto por Nelder e Wedderburn (1972) cresceu nas últimas décadas por abranger diversos tipos de distribuições para a variável resposta. Apesar do modelo ser um grande avanço na literatura estatística, ele possui suas limitações, como por exemplo: abranger apenas as distribuições da família exponencial, estrutura de linearidade, independência entre as observações da variável resposta, conforme Turkman e Silva (2000). O modelo linear clássico é um caso particular do MLG, em que sua variável resposta é modelada com uma distribuição normal, pertencente a família exponencial.

2.1 Definição

O modelo linear generalizado é composto por uma única variável resposta e suas variáveis explicativas. Elas são chamadas de componente aleatório e componente sistemático, respectivamente, e serão definidos a seguir juntamente com um outro elemento chamado de função de ligação.

Considere uma variável aleatória Y associada a um conjunto de variáveis explicativas x_1, x_2, \dots, x_k . Uma amostra aleatória de n observações é composta por (y_i, \mathbf{x}_i) , sendo que $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{ki})^T$ representando os valores das k covariáveis para a i -ésima unidade de observação da variável aleatória Y , $i = 1, \dots, n$.

1. O componente aleatório é formado por um conjunto das n realizações da variável resposta Y , Y_1, Y_2, \dots, Y_n , independentes, provenientes de uma mesma distribuição pertencente à família exponencial na forma canônica, com médias $\mu_1, \mu_2, \dots, \mu_n$. A função densidade de probabilidade (f.d.p) de Y_i é representada por

$$f(y_i; \theta_i, \phi) = \exp \left\{ \frac{1}{a_i(\phi)} [y_i \theta_i - b(\theta_i)] + c(y_i, \phi) \right\} \quad (2.2)$$

em que $b(\cdot)$ e $c(\cdot)$ são funções conhecidas, θ_i um parâmetro canônico e $\phi > 0$ um parâmetro de escala conhecido. Já $a_i(\phi) = \frac{\phi}{w_i}$, com w_i representando pesos a priori. A classe de distribuições pertencentes à família exponencial foi introduzida na

estatística por Fisher através de um estudo de propriedades de suficiência estatística, conforme Jørgensen e Labouriau (1992). Características dessa família são:

$$E(Y_i) = b'(\theta_i) = \mu_i$$

$$Var(Y_i) = a_i(\phi)b''(\theta_i) = a_i(\phi)V(\mu_i) = a_i(\phi)V_i$$

onde $b'(\theta_i) = \frac{\partial b(\theta_i)}{\partial \theta_i}$ e $V(\mu_i) = \frac{\partial \mu_i}{\partial \theta_i}$ representa a função de variância, função que depende unicamente da média, desempenhando assim um papel fundamental nos MLGs.

2. Componente sistemático: representa o conjunto de k variáveis explicativas do modelo, sendo $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{ki})^T$, onde $i = 1, 2, \dots, n$, que tem a característica de entrar no modelo na forma de uma estrutura linear, ou seja, soma linear dos seus efeitos, dado por

$$\eta_i = \sum_{j=1}^k x_{ij}\beta_j$$

ou na forma matricial representado por

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} \tag{2.3}$$

onde

$\mathbf{X}_{n \times p} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$ é a matriz de especificação do modelo (matriz de delineamento)

$\boldsymbol{\beta}_{p \times 1} = (\beta_1, \beta_2, \dots, \beta_p)^T$ é o vetor de parâmetros desconhecidos do modelo, cujo interesse é estimá-los

$\boldsymbol{\eta}_{n \times 1} = (\eta_1, \eta_2, \dots, \eta_n)^T$ é o vetor chamado de preditor linear

3. Função de ligação: representa a ligação entre os componentes aleatório e sistemático do modelo, relacionando a média ao preditor linear, portanto

$$\eta_i = g(\mu_i)$$

para $g(\cdot)$ uma função monótona e derivável. A Tabela 1 apresenta algumas funções de ligação canônicas.

Fonte: Demétrio (2002)

Tabela 1: Funções de ligação canônicas

Distribuição	Ligação canônica
Normal	Identidade: $\eta = \mu$
Poisson	Logarítmica: $\eta = \ln(\mu)$
Binomial	Logística: $\eta = \ln\left(\frac{\pi}{1-\pi}\right) = \ln\left(\frac{\mu}{m-\mu}\right)$
Gamma	Recíproca: $\eta = \frac{1}{\mu}$
Normal inversa	Recíproca ² : $\eta = \frac{1}{\mu^2}$

2.2 Estimação

Um das primeiras coisas a se fazer ao analisar um modelo de regressão é estimar o vetor de parâmetros $\boldsymbol{\beta}$, a fim de ajustar um modelo que represente a relação entre as variáveis. Os métodos utilizados para estimar esse vetor nos modelos de regressão linear (modelos clássicos) são os de máxima verossimilhança, mínimos quadrados, função geradora de momentos, entre outros. Já no MLG o método mais utilizado na literatura é o de máxima verossimilhança que será introduzido nessa seção.

Suponha y_1, y_2, \dots, y_n uma amostra aleatória independente de Y , tal que, $f(y_i; \theta_i; \phi)$ representa uma distribuição pertencente à família exponencial na forma canônica. Conforme Demétrio (2002) o logaritmo da função de verossimilhança é dado por:

$$\begin{aligned}
 \ell &= \ell(\boldsymbol{\theta}, \mathbf{y}) \\
 &= \sum_{i=1}^n \ln \left\{ \exp \left[\frac{1}{a_i(\phi)} [y_i \theta_i - b(\theta_i)] + c(y_i, \phi) \right] \right\} \\
 &= \sum_{i=1}^n \left\{ \frac{1}{a_i(\phi)} [y_i \theta_i - b(\theta_i)] + c(y_i, \phi) \right\}
 \end{aligned} \tag{2.4}$$

Para o desenvolvimento do processo de estimação, é importante citar que: $\ell = f(\theta_1, \theta_2, \dots, \theta_n / \mathbf{Y})$, $\theta_i = \int V_i^{-1} d\mu_i = q(\mu_i)$ e $\mu_i = g^{-1}(\eta_i)$, $\eta_i = \sum_{j=1}^k x_{ij} \beta_j$. Assim, ℓ depende de θ_i , que por sua vez depende de μ_i , que depende de η_i e que por último depende de β_j . Portanto, tem-se a seguinte estrutura:

$$\boldsymbol{\ell} \rightarrow \boldsymbol{\theta} \rightarrow \boldsymbol{\mu} \rightarrow \boldsymbol{\eta} \rightarrow \boldsymbol{\beta}$$

Derivando a equação (2.4) pela regra da cadeia, tem-se a seguinte expressão:

$$\begin{aligned}
U_j &= \sum_{i=1}^n \frac{\partial \ell_i}{\partial \beta_j} \\
U_j &= \sum_{i=1}^n \frac{\partial \ell_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j} \\
U_j &= \sum_{i=1}^n \frac{1}{a_i(\phi)} [y_i - b'(\theta_i)] \frac{1}{\frac{\partial \mu_i}{\partial \theta_i}} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j} \\
U_j &= \sum_{i=1}^n \frac{1}{a_i(\phi)} (y_i - \mu_i) \frac{1}{V(\mu_i)} \frac{\partial \mu_i}{\partial \eta_i} x_{ij}
\end{aligned} \tag{2.5}$$

Na maioria das vezes (2.5) não tem solução explícita, logo recorre a métodos iterativos para solucionar esse problema e obter as soluções desejadas. São exemplos de métodos iterativos o de Newton-Raphson, escore de Fisher, EM, entre outros. O método escore de Fisher fornece a seguinte solução

$$\hat{\boldsymbol{\beta}}^{(m+1)} = (\mathbf{X}^T \mathbf{W}^{(m)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{(m)} \mathbf{z}^{(m)} \tag{2.6}$$

sendo que \mathbf{X} é a matriz das variáveis explicativas do modelo, \mathbf{W} é a matriz diagonal dos pesos, onde cada elemento é representado por $W_i = \frac{w_i}{V(\mu_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)$ e $\mathbf{z}^{(m)} = \mathbf{X} \hat{\boldsymbol{\beta}}^{(m)} + \Delta^{(m)} (\mathbf{y} - \boldsymbol{\mu})^{(m)}$, representa o vetor da variável dependente ajustada no passo m do processo iterativo, com $\Delta = \text{diag} \left(\frac{\partial \eta_i}{\partial \mu_i} \right)$. Vale ressaltar que a expressão (2.6) não depende de ϕ .

Para obter a estimativa de $\boldsymbol{\beta}$ é necessário que se forneça ao algoritmo um valor inicial, chamado de chute inicial $\hat{\boldsymbol{\beta}}^{(0)}$. Quando a diferença de $|\hat{\boldsymbol{\beta}}^{(m-1)} - \hat{\boldsymbol{\beta}}^{(m)}|$ for muito pequena, $\hat{\boldsymbol{\beta}}^{(m)}$ é considerado a estimativa de máxima verossimilhança. O estimador obtido é um estimador consistente e eficiente de $\boldsymbol{\beta}$, conforme Paula (2010).

2.3 Comparação de modelos

Com o interesse em selecionar um modelo que melhor represente o problema em estudo, tem-se que vários critérios podem ser encontrados na literatura atual. Esses critérios são propostos com o intuito de obter o modelo mais adequado e com o menor número de covariáveis possíveis, dentre as diversas combinações de modelos. Vale ressaltar ainda que nem sempre os critérios levam à seleção de um mesmo modelo.

2.3.1 Critérios AIC e BIC

Esses dois critérios de avaliação do ajuste do modelo são altamente relacionados. O Critério de Informação de Akaike (AIC) foi proposto por Akaike (1974). Já o Critério de Informação Bayesiano (BIC) foi proposto por Schwarz (1978), e tem esse nome por conter argumentos Bayesianos em sua forma de avaliação.

O AIC não requer testes estatísticos e retorna um valor exato, onde este é uma medida para avaliar a qualidade do ajuste de um modelo estatístico. Quanto menor for esse valor melhor será o ajuste do modelo. Ele é baseado no logaritmo da função de verossimilhança, com a introdução de um fator de correção como modo de penalização da complexidade do modelo, conforme Turkman e Silva (2000). O critério de Akaike é representado por

$$AIC = -2\ell(\hat{\beta}) + 2p \quad (2.7)$$

sendo p o número de parâmetros no modelo e $\ell(\hat{\beta})$ representa o logaritmo da função de verossimilhança. Observa-se que o $\ell(\hat{\beta})$ depende do número de parâmetros do modelo, e a medida que aumenta o número de parâmetros no modelo o valor de $\ell(\hat{\beta})$ também cresce. Sendo assim, o principal objetivo é minimizar o valor do AIC para que possa obter um modelo com poucos parâmetros e que seja considerado como um dos melhores.

O BIC é o critério de informação Bayesiano bem similar ao AIC, pois também retorna um valor exato e é uma medida para avaliar o ajuste do modelo. Este critério é baseado na função de verossimilhança e também não requer testes estatísticos. Porém, ao analisar a estrutura de penalidade, ele é mais rigoroso que o AIC, pelo fato de que em alguns modelos o BIC é sensível ao aumento da verossimilhança. Portanto, ele penaliza mais fortemente que o critério de Akaike a introdução adicional de parâmetros, conforme Aiube (2007). O critério Bayesiano é representado por

$$BIC = -2\ell(\hat{\beta}) + p \ln n \quad (2.8)$$

onde p é o número de parâmetros no modelo, $\ell(\hat{\beta})$ representa o logaritmo da função de máxima verossimilhança e n o número de observações na amostra.

2.3.2 Deviance

Dando continuidade a seleção de modelos, tem-se agora o método de *deviance*. Este método é baseado em uma medida de discrepância do ajuste do modelo, o qual se mede a distância dos valores ajustados ($\hat{\mu}$) sob o modelo corrente ou modelo sob pesquisa, em

relação aos dados observados (y), isto é, modelo saturado.

O modelo corrente ou modelo sob pesquisa é definido como sendo um modelo com p parâmetros linearmente independentes, onde este modelo deve estar entre os modelos minimal e maximal. Como o nome já diz, o modelo minimal é aquele que será representado pelo mínimo de variáveis explanatórias, sendo estas influentes para aquele modelo. Já o maximal é o oposto, representa um modelo com o maior número de variáveis influentes. O modelo saturado ou completo é um modelo composto por todas as combinações das variáveis explanatórias, no qual este tem que conter n parâmetros, um para cada observação. O oposto desse é o modelo nulo, representado por um único parâmetro, sendo este igual a μ , comum a todos os dados. Conforme Demétrio (2002), a expressão da função *deviance* é dada por:

$$S_p = 2(\hat{\ell}_n - \hat{\ell}_p) \quad (2.9)$$

em que $\hat{\ell}_n$ representa o logaritmo da função de verossimilhança do modelo saturado e $\hat{\ell}_p$ do modelo corrente. É importante ressaltar que quanto menor for o S_p melhor será o ajuste do modelo.

2.4 Testes de Hipótese

Conforme Demétrio (2002), os testes de hipóteses são métodos inferenciais nos MLG e se baseiam fundamentalmente na teoria de máxima verossimilhança. A partir dessa teoria, serão citados aqui três testes para hipóteses relativas aos parâmetros β' s. Elas são deduzidas de distribuições assintóticas de funções adequadas das estimativas dos β' s. Essas estatísticas são:

- Razão de verossimilhança;
- Teste de Wald;
- Teste de Escore.

Elas são assintoticamente equivalentes e, sob H_0 e para ϕ conhecido, convergem para uma variável com distribuição χ_p^2 . Porém, a estatística da razão de verossimilhança é o critério que define um teste uniforme mais poderoso, conforme Demétrio (2002).

Conforme Paula (2010), para obter as três estatísticas referentes aos testes acima citados consideram-se as seguintes hipóteses simples:

$$H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}^0 \text{ contra } H_1 : \boldsymbol{\beta} \neq \boldsymbol{\beta}^0,$$

onde $\boldsymbol{\beta}^0$ é um vetor p -dimensional conhecido e assumindo também ϕ conhecido.

Estas hipóteses verificam a significância dos parâmetros em um modelo.

2.4.1 Razão de verossimilhança

O teste da razão de verossimilhança pode ser apresentado em duas formas: sendo uma a comparação da função de verossimilhança maximizada e a outra seria a comparação da função de desvio. No caso de hipóteses simples a estatística é representada por:

$$\xi_{RV} = 2\{\ell(\hat{\boldsymbol{\beta}}) - \ell(\boldsymbol{\beta}^0)\} = \phi\{D(\mathbf{y}, \boldsymbol{\mu}^0) - D(\mathbf{y}, \hat{\boldsymbol{\mu}})\}, \quad (2.10)$$

sendo $\ell(\hat{\boldsymbol{\beta}})$ o logaritmo da função de verossimilhança, $\ell(\boldsymbol{\beta}^0)$ logaritmo da função de verossimilhança sob a hipótese H_0 , $\boldsymbol{\mu}^0 = g^{-1}(\hat{\boldsymbol{\eta}}^0)$ e $\hat{\boldsymbol{\eta}}^0 = \mathbf{X}\hat{\boldsymbol{\beta}}^0$ também sob a hipótese H_0 .

Quando se tem amostras grandes e a estatística da razão de verossimilhança for maior que $\chi_{p,1-\alpha}^2$, rejeita-se H_0 a um nível de $100\alpha\%$.

2.4.2 Teste de Wald

Conforme Demétrio (2002), esse teste é baseado na distribuição assintótica de $\hat{\boldsymbol{\beta}}$ e é o mais utilizado quando se tem interesse em testar hipóteses relativas a um único coeficiente β_j . Assim, a estatística para o teste de Wald será dada por:

$$\xi_W = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0)^T [\hat{V}(\hat{\boldsymbol{\beta}})]^{-1} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0),$$

em que $\hat{V}(\hat{\boldsymbol{\beta}})^{-1}$ representa a matriz de informação de Fisher em $\hat{\boldsymbol{\beta}}$.

Um caso particular desse teste é quando se tem $p = 1$, pois o teste é equivalente ao quadrado do teste t usual

$$\xi_W = \frac{(\hat{\beta} - \beta^0)^2}{\hat{V}(\hat{\beta})}.$$

Quando se tem amostras grandes e a estatística de Wald for maior que $\chi_{p,1-\alpha}^2$, rejeita-se H_0 a um nível de $100\alpha\%$.

2.4.3 Teste Escore

Conforme Paula (2010), esse teste também é conhecido na literatura por Rao e é muito utilizado na área de Bioestatística. A estatística para esse teste quando $\mathbf{U}_\beta(\hat{\beta}) = \mathbf{0}$ é dada por:

$$\xi_{SR} = \mathbf{U}_\beta(\beta^0)^T \hat{V}(\hat{\beta}_0) \mathbf{U}_\beta(\beta^0),$$

onde $\hat{V}(\hat{\beta}_0)$ denota a variância assintótica de $\hat{\beta}$ que está sendo estimada sob H_0 .

Quando se tem amostras grandes e a estatística Escore for maior que $\chi_{p,1-\alpha}^2$, rejeita-se H_0 a um nível de $100\alpha\%$.

3 *MODELOS PARA VARIÁVEIS QUALITATIVAS*

Variáveis qualitativas se caracterizam por não poderem ser mensuradas numericamente e como o nome já diz se baseiam em classificações qualitativas ou categorias. Essas variáveis ainda podem se dividir em dois grupos, as nominais e ordinais que serão definidas posteriormente.

Os Modelos qualitativos são caracterizados por terem sua variável resposta como sendo do tipo qualitativa. São exemplos, os modelos logísticos nominais e os modelos logísticos para dados ordinais; esses últimos por sua vez abrangem os modelos logísticos acumulativos, modelo de chances proporcionais e modelos logísticos com categorias adjacentes, que são casos particulares do primeiro.

Um exemplo de dados medidos na escala ordinal são aqueles referentes à escala de Likert, muito utilizada na área de pesquisas de satisfação de clientes. Introduzida por Rensis Likert em 1932, tem como principal característica ser uma escala do tipo não comparativa, onde o entrevistado irá demonstrar seu grau de satisfação em um determinado item perguntado. Conforme Cunha (2007), um exemplo de escala Likert é aquela em que os níveis a ser escolhidos pelo entrevistado podem variar de

1. Discordo Totalmente
2. Discordo
3. Nem concordo nem discordo
4. Concordo
5. Concordo Totalmente.

Essa escala é uma das mais utilizadas na literatura, conhecida como sendo a de 5 níveis. Entretanto, essa escala pode variar até 11 níveis. Para o problema estudado neste trabalho, a variável a ser considerada é impressão global, originalmente representada por uma escala definida em 9 níveis e redefinida em uma escala de 5 níveis. Neste capítulo

serão abordados os modelos logísticos para dados nominais como forma de introdução aos modelos logísticos e em seguida o modelo de chances proporcionais, que será utilizado para a análise do problema aqui tratado.

3.1 Modelo logístico

A regressão logística é uma forma de modelagem estatística muito importante para dados de resposta categórica e é muito utilizada em diversos campos de aplicações, por exemplo: medicina, ciências, *marketing*, finanças, seguros, econometria, epidemiologia, entre outros. As categorias da variável resposta podem ser classificadas como sendo de natureza nominal ou ordinal. As variáveis de caráter nominal são aquelas representadas por itens que não envolvem uma ordem. A relação deste tipo de variável com uma ou mais covariáveis pode ser representada pelo modelo logístico nominal, que será abordado nesta seção.

3.1.1 Definição

Seja Y uma variável aleatória tal que $y = 1$ representa o evento sucesso e $y = 0$, o fracasso. Segundo Agresti (2002), a relação entre uma variável dependente (y) do tipo binária e suas variáveis independentes, pode ser representada pelo modelo de regressão logístico dado por:

$$\pi(\mathbf{x}) = \frac{\exp(\beta_0 + \boldsymbol{\beta}\mathbf{x})}{1 + \exp(\beta_0 + \boldsymbol{\beta}\mathbf{x})} \quad (3.1)$$

em que $\pi(\mathbf{x}) = P(Y = 1|\mathbf{X} = \mathbf{x}) = 1 - P(Y = 0|\mathbf{X} = \mathbf{x})$ e $0 < \pi(\mathbf{x}) < 1$, ou seja, $Y \sim \text{Bin}(1, \pi(\mathbf{x}))$ e $\boldsymbol{\beta}$ é o vetor de parâmetros do modelo associado a cada covariável.

Ainda conforme Agresti (2002) o *odds* de um evento que tem probabilidade p de ocorrer, é a probabilidade desse evento ocorrer dividido pela probabilidade dele não ocorrer, ou seja, o *odds* é dado por:

$$\text{odds} = \frac{p}{1 - p} \quad (3.2)$$

Considerando p como a probabilidade de sucesso, sendo $p = \pi(\mathbf{x})$, o *odds* pode ser interpretado como sendo $p = \pi(\mathbf{x})$ definido em (3.1). Assim, o *odds* é dado por:

$$\frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})} = \frac{\frac{\exp(\beta_0 + \boldsymbol{\beta}\mathbf{x})}{1 + \exp(\beta_0 + \boldsymbol{\beta}\mathbf{x})}}{1 - \frac{\exp(\beta_0 + \boldsymbol{\beta}\mathbf{x})}{1 + \exp(\beta_0 + \boldsymbol{\beta}\mathbf{x})}} = \exp(\beta_0 + \boldsymbol{\beta}\mathbf{x}). \quad (3.3)$$

Para linearizar o modelo, aplica-se o logaritmo nos dois lados da expressão (3.3), obtendo:

$$\text{logit}(\pi(\mathbf{x})) = \ln\left(\frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})}\right) = \ln(\exp(\beta_0 + \boldsymbol{\beta}\mathbf{x})) = \beta_0 + \boldsymbol{\beta}\mathbf{x}, \quad (3.4)$$

dando origem ao modelo $\text{logit}(\pi(\mathbf{x}))$, um caso particular do modelo linear generalizado, citado no capítulo anterior.

Modelar $\pi(\mathbf{x})$ com a função logística é equivalente a ajustar o modelo de regressão logístico onde a variável resposta é $\text{logit}(\pi(\mathbf{x}))$. Assim, a expressão do modelo de regressão logística é representado pela equação (3.4).

3.1.2 Estimação

Neste modelo a estimação dos parâmetros também é de suma importância. Foi mencionado anteriormente que o modelo de regressão logístico é um caso particular do modelo linear generalizado. Portanto, para obter a estimação dos parâmetros desse modelo, utiliza-se o mesmo procedimento da estimação do MLG.

Para estimar os parâmetros desse modelo, tem-se que, segundo Agresti (2002) a função de verossimilhança é dada por:

$$\mathbb{L}(\boldsymbol{\beta}) = \prod_{i=1}^n \pi(\mathbf{x})^{y_i} [1 - \pi(\mathbf{x})]^{1-y_i}, \text{ com } y_i \in \{0, 1\}. \quad (3.5)$$

Seja $\boldsymbol{\beta}$ o vetor de parâmetros relacionado com a probabilidade condicional $P(Y_i = 1|\mathbf{x}) = \pi(\mathbf{x})$. Para obter o estimador de $\boldsymbol{\beta}$ pelo método de máxima verossimilhança, com notação $\hat{\boldsymbol{\beta}}$, tem-se que derivando a equação (3.5) em relação aos $\boldsymbol{\beta}$'s, pode-se observar as seguintes equações

$$\begin{aligned} \sum_{i=1}^n (y_i - \pi(\mathbf{x})) &= 0 \\ \sum_{i=1}^n \mathbf{x}(y_i - \pi(\mathbf{x})) &= 0. \end{aligned} \quad (3.6)$$

Conforme Figueira (2006), a partir das equações de (3.6) pode-se verificar que $(k+1)$ representa o número de equações de verossimilhança, onde estas são obtidas ao diferenciar, em relação a $\boldsymbol{\beta}$, a função logaritmo de verossimilhança dada por

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^n [y_i \ln \pi(\mathbf{x}) + (1 - y_i) \ln(1 - \pi(\mathbf{x}))], \quad (3.7)$$

para cada um dos $k + 1$ coeficientes. A equação (3.7) é obtida a partir de (3.5) e do uso das propriedades de somatório e de logaritmo.

O sistema mostrado em (3.6) mostra equações não lineares, que não tem solução explícita, então, é de suma importância o uso de métodos iterativos. Como nos MLG's, um dos métodos que podem ser utilizados é o de Newton-Raphson. Com base nesse método pode-se observar que a equação (3.7) é uma função estritamente convexa, e assim a equação (3.6) admite um único ponto de máximo, ou seja, uma única solução.

Além disso, estima-se a probabilidade de cada categoria $\pi(\mathbf{x})$. Manipulando a equação (3.4) tem-se a seguinte expressão

$$\begin{aligned} \ln\left(\frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})}\right) &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} \\ \frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})} &= e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}} \\ \pi(\mathbf{x}) &= \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}}} = \frac{e^\theta}{1 + e^\theta}, \end{aligned} \quad (3.8)$$

sendo θ um parâmetro canônico do modelo representado por $\beta_0 + \sum_{j=1}^k \beta_j x_{ij}$. A estimativa dessa probabilidade é representada por:

$$\hat{\pi}(\mathbf{x}) = \frac{e^{\hat{\theta}}}{1 + e^{\hat{\theta}}} \quad (3.9)$$

3.1.3 Odds Ratio

Além das estimativas dos parâmetros um outro objetivo no modelo logístico pode ser o de obter o valor do *odds ratio* das variáveis em estudo. O *odds ratio* é um método muito utilizado em tabelas de contingência, sendo de grande importância na área da saúde. Seu objetivo é estabelecer a relação de dois *odds* (ou duas chances) sobre uma determinada característica em comum. De acordo com Agresti (2002) ela é representada por

$$\begin{aligned} OR &= \frac{\pi_{11}/(1 - \pi_{11})}{\pi_{22}/(1 - \pi_{22})} = \frac{\pi_{11}/\pi_{21}}{\pi_{12}/\pi_{22}} = \frac{\pi_{11}\pi_{22}}{\pi_{21}\pi_{12}} \\ &= \frac{P(Y = 1|X = 1)P(Y = 2|X = 2)}{P(Y = 2|X = 1)P(Y = 1|X = 2)}, \end{aligned} \quad (3.10)$$

sendo π_{11} : probabilidade do evento $(Y = 1|X = 1)$ ocorrer, π_{12} : probabilidade do evento $(Y = 1|X = 2)$ ocorrer, e que $\pi_{11} + \pi_{21} = 1$ e $\pi_{12} + \pi_{22} = 1$.

Exemplo: Um departamento da área da saúde elaborou uma pesquisa. Seu objetivo é comparar a chance de mulheres com câncer que usam anticoncepcional (ACO) com as que não usam. Essa pesquisa pode ser representada pela seguinte tabela

Tabela 2: Mulheres com câncer de mama

Câncer	Usaram ACO	Não usaram ACO
Sim	273	716
Não	2641	7260
Total	2914	7976

Fonte: Dados fictícios

Onde:

Evento A: Mulher com câncer dado que usou anticoncepcional

Evento B: Mulher com câncer dado que não usou anticoncepcional

Sendo $\hat{\pi}_{11} = \frac{273}{2914}$ a probabilidade do evento A ocorrer e $\hat{\pi}_{12} = \frac{716}{7976}$ a probabilidade B ocorrer. O *odds ratio* do evento A e B será representado por:

$$OR_{A|B} = \frac{\frac{\frac{273}{2914}}{1 - \frac{273}{2914}}}{\frac{\frac{716}{7976}}{1 - \frac{716}{7976}}} = 1,047$$

Portanto estima-se que mulheres que usaram anticoncepcional tem 4,7% mais chance de desenvolver câncer de mama do que as mulheres que não usaram anticoncepcional.

Trazendo essa definição de *odds ratio* para o modelo de regressão logístico, é preciso ter obtido antes os valores das estimativas dos parâmetros, pois o *odds ratio* nesse modelo depende dos parâmetros.

Substituindo a equação (3.9) na (3.10), tem-se a seguinte expressão para estimativa do *odds ratio* no modelo para a categoria j

$$\hat{OR}_j = \frac{\left(\frac{\exp\{\beta_0 + \sum_{i=1}^k \beta_j x_{ij}\}}{1 + \exp\{\beta_0 + \sum_{i=1}^k \beta_j x_{ij}\}} \right)}{\left(\frac{\exp\{\beta_0 + \sum_{i=1}^k \beta_j x_{ij}\}}{1 + \exp\{\beta_0 + \sum_{i=1}^k \beta_j x_{ij}\}} \right)} = \exp^{\hat{\beta}_j}, \quad (3.11)$$

podendo ser utilizada para modelos de variáveis dicotômicas (duas categorias) ou politômicas (mais de duas categorias), porém ao modelar este último caso é necessário dicotomizá-las.

3.1.3.1 Intervalo de confiança para o *odds ratio*

Dando continuidade a esse processo inferencial sobre o *odds ratio* deve-se obter a estimativa intervalar para esta quantidade, uma vez que a estimação pontual não leva em conta a variabilidade amostral. No intervalo de confiança o processo depende dos valores estimados dos β 's e seus respectivos erros padrões. Estes valores estão apenas relacionados com a categoria j que se deseja estudar. Logo, para obter esse intervalo de confiança para *odds ratio* da equação (3.11), tem-se que esse processo será representado pela seguinte expressão:

$$IC_{1-\alpha}(OR_j) = \exp[\beta_j \pm z_{\alpha/2} EP(\beta_j)] \quad (3.12)$$

sendo $z_{\frac{\alpha}{2}}$ o quantil da normal de ordem $\alpha/2$ de significância e $EP(\beta_j)$ o erro padrão do coeficiente.

3.1.3.2 Teste de hipótese para *odds ratio*

Para este processo é importante lembrar que o valor do *odds ratio* depende do parâmetro β respectivo a categoria de interesse, podendo este valor ser observado a partir da equação (3.11). Por exemplo, o teste de hipótese para o *odds ratio* = 1 equivale testar o $\beta = 0$. Sendo assim, o correto é utilizar o teste dos parâmetros β para tirar conclusões sobre o *odds ratio*.

Para testar os parâmetros β 's pode ser utilizado o teste de razão de verossimilhança, pois foi mencionado no capítulo anterior que ele é mais eficiente em relação aos outros testes. A hipótese a ser testada é dada por

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

Esse teste tem a pressuposição de que $\hat{\beta} \sim \mathcal{N}(\beta, Var(\hat{\beta}))$, isto é $\frac{\hat{\beta} - \beta}{EP(\hat{\beta})} \sim \mathcal{N}(0, I)$. Logo, a estatística desse teste é representada pela equação (2.10).

3.2 Modelo de chances proporcionais (MCP)

O modelo de chances proporcionais surgiu na literatura como sendo uma generalização dos modelos logísticos acumulativos. Agresti (2002) descreve o modelo logístico acumu-

lativo como uma forma de modelagem de respostas categóricas ordenadas, onde no *logit cumulative* obtem-se probabilidades acumuladas, sendo expressas por

$$\gamma_h(\mathbf{x}) = \pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \dots + \pi_h(\mathbf{x})$$

sendo h a categoria de resposta contida no conjunto $1, 2, \dots, H$ e γ_h a função acumulada até a categoria h . Dobson e Barnett (1945) descrevem os modelos acumulativos como tendo um intercepto comum para todas as categorias. Esse modelo é dado por

$$\log \left(\frac{\pi_1(x) + \pi_2(x) + \dots + \pi_h(x)}{\pi_{h+1}(x) + \pi_{h+2}(x) + \dots + \pi_H(x)} \right) = \mathbf{x}^T \boldsymbol{\beta}_j. \quad (3.13)$$

Os modelos de chances proporcionais tem como característica calcular um valor de intercepto para cada categoria, ou seja, com base no número de categorias será calculado o número de interceptos e estes terão um peso proporcionalmente diferente para cada uma das categorias. Nesse modelo também é importante ressaltar que esse intercepto depende unicamente das categorias e não das variáveis explicativas.

Segundo McCullagh (1980) os modelos de chances proporcionais são mais úteis na prática por causa de sua simplicidade e interpretação. Como dito anteriormente, esses modelos são baseados em probabilidades acumuladas, sendo possível estimar as probabilidades marginais ocorridas em cada categoria. Uma suposição desse modelo é que o coeficiente de inclinação é o mesmo para cada categoria h e uma curiosidade é que a chance de cada categoria ocorrer é proporcional em cada categoria, por isso, o modelo é classificado como de chances proporcionais. A Figura 1 mostra que nesse modelo é possível verificar que o efeito das variáveis explicativas sobre a escala logarítmica das chances acumuladas é sempre igual em todas as H categorias, contudo na escala do log do *odds* a probabilidade das categorias são representadas por linhas paralelas, conforme Dobson e Barnett (2008).

Para mais detalhes sobre esse modelo, verificar McCullagh (1980), Dobson e Barnett (2008), Okura (2008) e Marôco (2011).

3.2.1 Definição

Conforme Lara (2007), suponha uma amostra aleatória de N indivíduos, onde a variável resposta é do tipo ordinal, isto é, suas categorias podem assumir valores no conjunto $h = \{1, 2, 3, \dots, H\}$, tais que, $1 < 2 < \dots < H$. A resposta para o i -ésimo indivíduo é denotada pelo vetor $\mathbf{y}_i = (I_{i1}, I_{i2}, \dots, I_{iH})^T$, em que I_{ih} 's representam as

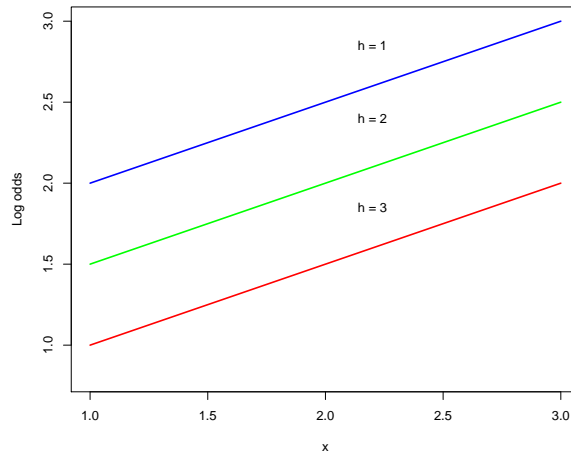


Figura 1: Modelo de chances proporcionais para a escala do log *odds*

variáveis indicadoras para as categorias de resposta, isto é, $I_{ih} = 1$ se o indivíduo optou por esta h -ésima categoria e $I_{ih} = 0$, caso contrário. Seja $\mathbf{x} = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ o vetor de dimensão p . O $\pi_h(\mathbf{x})$ representa a probabilidade marginal de ocorrência de I_{ih} . As probabilidades acumuladas são representadas por:

$$\begin{aligned}\gamma_1(\mathbf{x}) &= \pi_1(\mathbf{x}) \\ \gamma_2(\mathbf{x}) &= \pi_1(\mathbf{x}) + \pi_2(\mathbf{x})\end{aligned}$$

Logo:

$$\gamma_H(\mathbf{x}) = \pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \dots + \pi_H(\mathbf{x}) \quad (3.14)$$

Portanto, a probabilidade acumulada do indivíduo para a categoria H , será representado por:

$$\gamma_H(\mathbf{x}) = 1$$

pois representa a soma das probabilidades até H .

Para ajustar um modelo de chances proporcionais a um conjunto de dados, tem-se que para cada categoria, o modelo a ser ajustado será representado por:

$$\ln \left[\frac{\gamma_h(\mathbf{x})}{1 - \gamma_h(\mathbf{x})} \right] = \beta_{0h} + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_k) = \beta_{0h} + \boldsymbol{\beta}' \mathbf{x} \quad (3.15)$$

em que β_{0h} é o intercepto para a categoria h e \mathbf{x} representa o vetor de valores das variáveis explicativas. Ao ajustar um modelo do tipo (3.15) em que a variável dependente é do tipo binomial (dicotômica), podendo assumir apenas dois valores, tem-se que o modelo

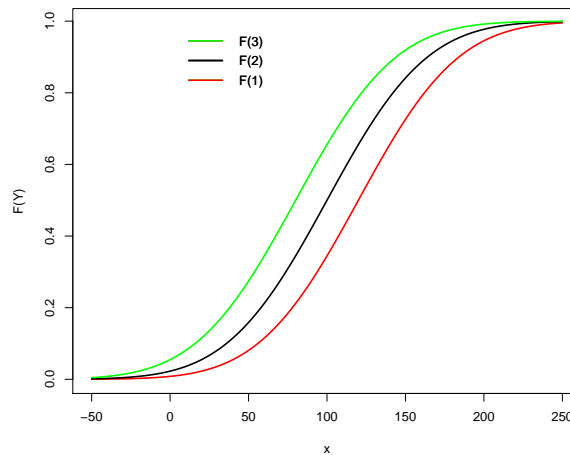
proposto acima reduz-se ao modelo de regressão logística para dados binários.

Como em todos os modelos, é possível verificar a influência dos valores de $\boldsymbol{\beta}$ sobre os valores da variável Y . No modelo (3.15), quando tem-se $\beta_i > 0$ e os valores assumidos por \mathbf{x} aumentam, pode-se observar que a probabilidade da variável dependente assumir valores nas categorias inferiores ou iguais a h também aumentam, isto é, Y tende a ser menor para valores alto de x . Então, para obter maiores valores de Y quando os valores de \mathbf{x} forem elevados e o $\beta_i > 0$ é preciso reescrever o modelo (3.15), onde o $\boldsymbol{\beta}$ será substituído por $-\boldsymbol{\beta}$. O modelo reparametrizado é definido por:

$$\ln \left[\frac{\gamma_h(\mathbf{x})}{1 - \gamma_h(\mathbf{x})} \right] = \beta_{0h} - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k) = \beta_{0h} - \boldsymbol{\beta}' \mathbf{x}. \quad (3.16)$$

Esse comportamento pode ser observado com base na Figura 2, onde os valores de X e Y variam, $F(Y)$ é igual à probabilidade acumulada de $P(Y \leq h)$ e que também é igual a $\gamma_h(\mathbf{x})$.

Figura 2: Comportamento da função de probabilidade do modelo de chances proporcionais



3.2.2 Estimação

O método a ser apresentado nesse tipo de modelo não será muito distinto do método apresentado no capítulo anterior. Como nos modelos anteriores tem-se a finalidade de estimar o vetor de parâmetros $\boldsymbol{\beta}$. Além disso, é de interesse estimar as probabilidades associadas a cada categoria na escala da variável resposta. Será utilizada a teoria da máxima verossimilhança por meio do método iterativo de *Newton - Raphson*. Os resultados obtidos por esse método geram expressões não lineares, como mencionado nos casos

anteriores. Conforme Agresti (2002), a função de verossimilhança é representada por

$$\begin{aligned} L(\boldsymbol{\gamma}, \boldsymbol{\beta}) &= \prod_{i=1}^n \left[\prod_{h=1}^H (\gamma_h(x_i) - \gamma_{h-1}(x_i))^{y_{ih}} \right] = \\ &= \prod_{i=1}^n \left[\prod_{h=1}^H \left(\frac{e^{(\beta_{0h} - x_i \boldsymbol{\beta})}}{1 + e^{(\beta_{0h} - x_i \boldsymbol{\beta})}} - \frac{e^{(\beta_{0(h-1)} - x_i \boldsymbol{\beta})}}{1 + e^{(\beta_{0(h-1)} - x_i \boldsymbol{\beta})}} \right)^{y_{ih}} \right] \end{aligned} \quad (3.17)$$

Uma vez obtidas as estimativas dos parâmetros, pode-se estimar as probabilidades acumuladas por meio de:

$$\hat{\gamma}_h(\mathbf{x}) = \frac{\exp(\hat{\beta}_{0h} - \hat{\boldsymbol{\beta}}\mathbf{x})}{1 + \exp(\hat{\beta}_{0h} - \hat{\boldsymbol{\beta}}\mathbf{x})} \quad (3.18)$$

sendo $h = 1, 2, \dots, H - 1$, pois por definição a probabilidade acumulada até a última categoria de resposta é igual a 1. Modelando a probabilidade acumulada pode-se obter os valores das probabilidades marginais de cada categoria, expressada na equação a seguir

$$\hat{\pi}_h(\mathbf{x}) = \hat{\gamma}_h(\mathbf{x}) - \hat{\gamma}_{(h-1)}(\mathbf{x}). \quad (3.19)$$

Por exemplo, calcular a probabilidade marginal de um indivíduo estar na categoria 4, será representado por:

$$\hat{\pi}_4(\mathbf{x}) = \hat{\gamma}_4(\mathbf{x}) - \hat{\gamma}_3(\mathbf{x})$$

Os modelos abordados neste capítulo são de efeitos fixos, a seguir consideram-se a inclusão de efeitos aleatórios, ou seja, modelos mistos.

4 *MODELOS MISTOS*

Os modelos mistos tem como principal característica a inclusão de um efeito aleatório, diferente do modelo clássico, que leva em consideração apenas o efeito fixo. Pelo fato de conter efeito aleatório esse tipo de modelo é chamado de modelos de efeitos aleatórios ou modelos de efeitos mistos ou apenas modelos mistos.

Os modelos mistos são conhecidos na literatura como sendo indicados para representar o comportamento de dados que apresentam dependência entre si, ou seja, uma observação depende estocasticamente da outra, conforme Costa (2010). Nessa metodologia é possível obter as correlações entre e intra grupos, além de estudar e manusear a estrutura de covariância, sendo esta de suma importância nos modelos de dados agrupados. Uma outra característica desse modelo é o fato de poder mensurar o comportamento individual e médio das observações, diferentes dos modelos usuais como o clássico que mensura apenas o comportamento médio.

Essa metodologia começou a ser muito utilizada na avaliação genética de bovinos. Com o decorrer dos anos foi sendo mais descrita na literatura. Atualmente, vem sendo bastante utilizada nas diversas áreas de estudo, principalmente na biológica.

Nesse capítulo serão propostos duas classes de modelos mistos. Os modelos lineares mistos clássicos e os modelos lineares generalizados mistos.

4.1 *Modelo linear misto clássico*

Essa classe de modelos mistos é uma extensão dos modelos lineares clássicos, pois levam em consideração o conjunto de covariáveis fixas (efeitos fixos) e a inclusão de um efeito aleatório, constituindo o modelo linear misto.

Conforme Demidenko (2004) o modelo linear misto foi introduzido por Laird e Ware (1982), sendo representado por:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \mathbf{e}_i, \quad i = 1, \dots, n \quad (4.1)$$

onde \mathbf{y}_i é o vetor ($n_i \times 1$) da variável resposta associada à i -ésima unidade amostral, $\boldsymbol{\beta}$ é o vetor ($p \times 1$) de parâmetros desconhecidos relacionados com os efeitos fixos, \mathbf{X}_i é a matriz ($n_i \times p$) de constantes conhecidas relacionadas aos efeitos fixos, \mathbf{Z}_i é a matriz ($n_i \times q$) de constantes conhecidas relacionada ao efeito aleatório, \mathbf{u}_i é o vetor ($q \times 1$) de parâmetros desconhecidos denominados de efeito aleatório associado à i -ésima unidade amostral e \mathbf{e}_i é o vetor ($n_i \times 1$) de erros aleatórios, que representam a variabilidade residual das observações realizadas na i -ésima unidade amostral. Adota-se que:

- $\mathbf{e}_i \stackrel{\text{ind}}{\sim} \mathcal{N}_{n_i}(\mathbf{0}, \sigma^2 \mathbf{I})$, sendo $i = 1, 2, \dots, n$, onde $\mathbf{I} = \mathbf{I}_{n_i}$ é uma matriz ($n_i \times n_i$) identidade;
- $\mathbf{u}_i \stackrel{\text{iid}}{\sim} \mathcal{N}_q(\mathbf{0}, \sigma^2 \mathbf{D})$, sendo $i = 1, 2, \dots, n$, onde \mathbf{D} é uma matriz ($q \times q$) definida positiva;
- \mathbf{e}_i e \mathbf{u}_i são independentes $\forall i$.

Conforme a expressão (4.1), tem-se que:

$$E(\mathbf{y}_i) = \mathbf{X}_i \boldsymbol{\beta}$$

$$Var(\mathbf{y}_i) = \mathbf{V} = \sigma^2 \mathbf{Z}_i \mathbf{D} \mathbf{Z}_i^T + \sigma^2 \mathbf{I} = \sigma^2 (\mathbf{Z}_i \mathbf{D} \mathbf{Z}_i^T + \mathbf{I}).$$

A matriz $\sigma^2 \mathbf{D}$ pode assumir diversas estruturas de covariância, onde essa estrutura irá depender da particularidade do estudo proposto pelo pesquisador.

4.1.1 Estimação

Na atualidade existem diversos métodos estatísticos que proporcionam a estimação dos parâmetros em um modelo de regressão. Segundo Nobre (2004) destacam-se os métodos bayesiano, máxima verossimilhança, máxima verossimilhança restrita e mínimos quadrados, sendo estes métodos equivalentes, supondo σ^2 e \mathbf{D} conhecidos, desde que no método bayesiano seja atribuída uma distribuição a priori não informativa para \mathbf{u} .

Deseja-se estimar o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma^2, \text{vet}(\mathbf{D}))$, em que $\text{vet}(\mathbf{D})$ representa o vetor formado pelos elementos distintos da matriz \mathbf{D} . A partir das suposições feitas anteriormente, pode-se observar que o espaço paramétrico que deve estimar o vetor de $\boldsymbol{\theta}$ é representado por:

$$\Theta = \{\boldsymbol{\theta} : \boldsymbol{\beta} \in R^m, \sigma^2 > 0, \mathbf{D} \text{ é uma matriz definida positiva}\}$$

Ainda segundo Nobre (2004), Harville (1976) utilizou o teorema de Gauss-Markov para obter o melhor estimador linear não viciado (BLUE) para β e o melhor preditor linear não viciado (BLUP) para u

Quando os parâmetros ligados à covariância são desconhecidos existem divergências quanto ao procedimento de estimação, conforme Hilden-Minton (1995). O método mais utilizado para estimar o vetor de parâmetros ligado a covariância é o de máxima verossimilhança restrita (MVR), visto que foi constatado que esse método retorna estimativas menos viesadas do que as estimativas dos outros métodos.

Conforme Davis (2002) o método de máxima verossimilhança restrita (MVR) tem menos viés com relação aos outros métodos por que o MVR está associada aos “contrastes dos erros” e não às observações reais.

Com o intuito de obter os estimadores utilizando o método MVR, é necessário isolar a parcela da função de verossimilhança que está associada apenas à estrutura de covariância do modelo. Esse isolamento pode ser feito ao considerarmos a transformação linear $Q^T y$, em que:

$$Q = \sigma^2(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}\mathbf{V}^{-1})$$

$$Q^T\mathbf{X} = \mathbf{0}_{(n \times p)} \quad \text{e} \quad Q^T y \sim \mathcal{N}(\mathbf{0}_n, Q^T\mathbf{V}^{-1}Q)$$

A partir dessas definições, tem-se que o logaritmo da função de MVR é dado por:

$$l_R(\beta) = -\frac{1}{2} \sum_{j=1}^k \{ \ln|\mathbf{X}^T\mathbf{V}\mathbf{X}| + \ln|\mathbf{V}| + (\mathbf{y} - \mathbf{X}\beta)^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta) \} \quad (4.2)$$

Conforme Demidenko (2004) a expressão (4.2) é equivalente ao logaritmo da função de verossimilhança dos resíduos, sendo $\hat{e} = \mathbf{y} - \mathbf{X}\hat{\beta}$. Muitas vezes o logaritmo da função de verossimilhança dos resíduos é chamada de método de máxima verossimilhança residual, e representada por:

$$l_R(\hat{e}) = -\frac{1}{2} \sum_{j=1}^k \{ \ln|\mathbf{X}^T\mathbf{V}\mathbf{X}| + \ln|\mathbf{V}| + \hat{e}^T\hat{e} \} \quad (4.3)$$

Geralmente usam-se métodos iterativos para estimar os parâmetros de covariância, pois essas funções são não lineares e a partir disso métodos como o de Newton-Raphson e algoritmo EM precisam ser utilizados. Portanto a expressão (4.3) é mais conveniente

computacionalmente na hora de obter esses estimadores.

Vale ressaltar que o método de máxima verossimilhança restrita e o máxima verossimilhança são assintoticamente equivalentes.

4.2 Modelo linear generalizado misto (MLGM)

Essa classe de modelos é uma extensão dos modelos lineares mistos e tem a mesma estrutura dos modelos lineares generalizados propostos por Nelder e Wedderburn (1972). Contudo, os modelos lineares generalizados mistos incluem um vetor de efeitos aleatórios \mathbf{u} no preditor linear, estrutura semelhante aos modelos lineares mistos.

Conforme McCulloch e Searle (2001) para obter o modelo linear generalizado misto é preciso considerar “distribuição condicional de \mathbf{Y} dado \mathbf{u} ”, onde \mathbf{Y} representa o vetor de respostas, sendo estas condicionalmente independentes com densidade pertencente à família exponencial na forma canônica. Logo tem-se que:

$$Y_i|\mathbf{u} \stackrel{\text{ind}}{\sim} f_{Y_i|\mathbf{u}}(y_i|\mathbf{u})$$

com

$$f_{Y_i|\mathbf{u}}(y_i|\mathbf{u}) = \exp \left\{ \frac{w_i}{\phi} [y_i\theta_i - b(\theta_i)] + c(y_i; \phi) \right\}. \quad (4.4)$$

Note que a relação entre a média de Y_i e a função $b(\theta_i)$ é semelhante àquela encontrada no MLG, sendo válida também aqui. Portanto a $E(Y_i|\mathbf{u}) = \mu_i = b'(\theta) = \frac{\partial b(\theta_i)}{\partial \theta_i}$.

Para este modelo a função de ligação será representada por:

$$g(\mu_i) = \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{z}_i^T \mathbf{u}, \quad (4.5)$$

onde \mathbf{x}_i representa a i -ésima linha da matriz do modelo para os efeitos fixos; $\boldsymbol{\beta}$ representa o vetor ($p \times 1$) de parâmetros com efeitos fixos; \mathbf{z}_i^T representa a i -ésima linha da matriz do modelo para os efeitos aleatórios e por último o \mathbf{u} que representa o vetor ($q \times 1$) de efeitos aleatórios.

Conforme Costa (2010) é importante atribuir uma distribuição aos efeitos aleatórios para completar as especificações do modelo. A distribuição será representada por:

$$\mathbf{u} \sim f_U(\mathbf{u}). \quad (4.6)$$

A distribuição mais usual que se atribui para os efeitos aleatórios é a distribuição normal, ou seja, $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{G})$. Abordagem para os efeitos aleatórios não normais pode ser vista em Lee e Nelder (1996), que utilizam distribuições conjugadas para esses efeitos.

Com base no modelo condicional (4.4), tem-se que a média da distribuição marginal de \mathbf{Y} é representada por:

$$E(Y_i) = E[E(Y_i|\mathbf{u})] = E(\mu_i) = E[g^{-1}(\mathbf{x}_i^T \beta + \mathbf{z}_i^T \mathbf{u})].$$

Em geral, a expressão acima não pode ser simplificada, por conta da presença de funções não-lineares em g^{-1} .

Já variância marginal de \mathbf{Y} é representada por:

$$\begin{aligned} V(Y_i) &= V[E(Y_i|\mathbf{u})] + E[V(Y_i|\mathbf{u})] \\ &= V(\mu_i) + E(\phi V(\mu_i)) \\ &= V[g^{-1}(\mathbf{x}_i^T \beta + \mathbf{z}_i^T \mathbf{u})] + E\{\phi V[g^{-1}(\mathbf{x}_i^T \beta + \mathbf{z}_i^T \mathbf{u})]\}. \end{aligned}$$

A expressão acima só poderá ser simplificada se houver suposições específicas sobre a forma de $g(\cdot)$ e/ou da distribuição condicional de \mathbf{Y} .

Ao usar um efeito aleatório no modelo é possível observar que ele introduz uma correção entre as observações que tem algum efeito em comum. Ao assumir independência condicional dos elementos de \mathbf{Y} , tem-se que a covariância será representada por:

$$\begin{aligned} Cov(Y_i, Y_j) &= Cov[E(Y_i|\mathbf{u}), E(Y_j|\mathbf{u})] + E[Cov(Y_i, Y_j|\mathbf{u})] \\ &= Cov(\mu_i, \mu_j) + E(0) \\ &= Cov[g^{-1}(\mathbf{x}_i^T \beta + \mathbf{z}_i^T \mathbf{u}), g^{-1}(\mathbf{x}_k^T \beta + \mathbf{z}_i^T \mathbf{u})] \end{aligned}$$

4.3 MCP com efeito aleatório

Com o objetivo de modelar a possível dependência existente entre as observações repetidas em uma mesma unidade experimental, este trabalho considera a inclusão do efeito aleatório no modelo de chances proporcionais. De acordo com Christensen (2012), o modelo que pode ser considerado é dado por:

$$\ln \left[\frac{\gamma_{ih}(\mathbf{x})}{1 - \gamma_{ih}(\mathbf{x})} \right] = \beta_{0h} - \beta' \mathbf{x}_i - u_i, \quad (4.7)$$

em que o termo u_i corresponde ao efeito aleatório da i -ésima unidade experimental, supondo que $u_i \sim \mathcal{N}(0, \sigma_u^2)$.

4.3.1 Estimação

Na atualidade existem diversos métodos para estimar os parâmetros do modelo (4.4). Nesta seção abordamos o de máxima verossimilhança sugerido por McCulloch e Searle (2001).

Como \mathbf{Y} segue a distribuição apresentada em (4.4) e o efeito aleatório a distribuição em (4.6), tem-se que a função de verossimilhança será representada por:

$$L = \int_{\mathbb{R}} \prod_{i=1}^n f_{Y_i|\mathbf{u}}(y_i|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \quad (4.8)$$

onde a integração é feita em relação a distribuição de \mathbf{u} que tem dimensões ($q \times 1$).

Pode-se observar que para maximizar a equação (4.8) é de grande necessidade utilizar métodos numéricos de integração. Na literatura existem diversas técnicas, como por exemplo os métodos bayesianos, a quadratura de Gauss-Hermite, métodos de quase-verossimilhança penalizada, entre outros.

Para dar continuidade à estimação dos parâmetros nos MLGMs é importante obter as equações de verossimilhança para os efeitos fixos e para os efeitos aleatórios. As respectivas equações estão apresentadas a seguir.

4.3.1.1 Função de verossimilhança para o efeito fixo

Conforme McCulloch e Searle (2001) a solução para as equações de verossimilhança são numericamente difíceis. Contudo existe uma forma mais simples de reescreve-las. A partir da equação (4.8) tem-se que:

$$\ell = \log \int_{\mathbb{R}^q} f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} = \log f_{\mathbf{Y}}(\mathbf{y}). \quad (4.9)$$

Derivando a equação (4.9) em relação ao efeito fixo β , tem-se que:

$$\begin{aligned}\frac{\partial \ell}{\partial \beta} &= \frac{\partial}{\partial \beta} \left[\log \int_{\mathbb{R}^q} f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \right] \\ &= \frac{1}{\int_{\mathbb{R}^q} f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}} \left[\frac{\partial}{\partial \beta} \int_{\mathbb{R}^q} f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \right] \\ &= \frac{1}{f_{\mathbf{Y}}(\mathbf{y})} \int_{\mathbb{R}^q} \left[\frac{\partial}{\partial \beta} f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) \right] f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}\end{aligned}\quad (4.10)$$

como $f_{\mathbf{U}}(\mathbf{u})$ não depende de β . Assim, dando continuidade tem-se que:

$$\begin{aligned}\frac{\partial}{\partial \beta} f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) &= \left[\frac{1}{f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u})} \frac{\partial}{\partial \beta} f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) \right] f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) \\ &= \left[\frac{\partial}{\partial \beta} \log f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) \right] f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}).\end{aligned}\quad (4.11)$$

Substituindo a equação (4.11) em (4.10), tem-se que a nova equação será dada por:

$$\begin{aligned}\frac{\partial \ell}{\partial \beta} &= \frac{1}{f_{\mathbf{Y}}(\mathbf{y})} \int_{\mathbb{R}^q} \left[\frac{\partial}{\partial \beta} \log f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) \right] f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \\ &= \int_{\mathbb{R}^q} \left[\frac{\partial}{\partial \beta} \log f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) \right] \frac{f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u}}{f_{\mathbf{Y}}(\mathbf{y})} \\ &= \int_{\mathbb{R}^q} \left[\frac{\partial}{\partial \beta} \log f_{\mathbf{Y}|\mathbf{u}}(\mathbf{y}|\mathbf{u}) \right] f_{\mathbf{U}|\mathbf{Y}}(\mathbf{u}|\mathbf{y}) d\mathbf{u}.\end{aligned}\quad (4.12)$$

Lembrando que a derivada da função de log-verossimilhança de um MLG é representada por:

$$\frac{\partial \ell}{\partial \beta} = \mathbf{X}^T \mathbf{W} \Delta (\mathbf{y} - \boldsymbol{\mu})$$

em que $\mathbf{W} = \text{diag}\{w_i\} = \text{diag}\left\{\frac{1}{\phi V_i} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2\right\}$ e $\Delta = \text{diag}\left\{\frac{\partial \eta_i}{\partial \mu_i}\right\}$.

Reescrevendo a equação (4.12), tem-se que:

$$\frac{\partial \ell}{\partial \beta} = \int_{\mathbb{R}^q} \mathbf{X}^T \mathbf{W}^* (\mathbf{y} - \boldsymbol{\mu}) f_{\mathbf{U}|\mathbf{Y}}(\mathbf{u}|\mathbf{y}) d\mathbf{u}$$

onde $\mathbf{W}^* = \text{diag}\left[\phi V_i \frac{\partial \eta_i}{\partial \mu_i}\right]^{-1}$. Então,

$$\frac{\partial \ell}{\partial \beta} = \mathbf{X}^T y E[\mathbf{W}^* | \mathbf{y}] - \mathbf{X}^T E[\mathbf{W}^* \boldsymbol{\mu} | \mathbf{y}]$$

Assim, a equação de verossimilhança para o parâmetro β será dada por:

$$\mathbf{X}^T y E[\mathbf{W}^* | \mathbf{y}] = \mathbf{X}^T E[\mathbf{W}^* \boldsymbol{\mu} | \mathbf{y}].$$

4.3.1.2 Função de verossimilhança para o efeito aleatório

Ao derivar o logaritmo da função de verossimilhança em relação a $\varphi = \sigma^2$, tem-se que o resultado para a equação da verossimilhança do parâmetro do efeito aleatório é bem similar ao encontrado em (4.12). Então, utilizando a equação (4.9) tem-se que:

$$\begin{aligned}
\frac{\partial \ell}{\partial \varphi} &= \frac{1}{f_{\mathbf{Y}}(\mathbf{y})} \frac{\partial f_{\mathbf{Y}}(\mathbf{y})}{\partial \varphi} \\
&= \frac{1}{f_{\mathbf{Y}}(\mathbf{y})} \frac{\partial}{\partial \varphi} \left[\int_{\mathbb{R}^q} f_{\mathbf{Y}|\mathbf{U}}(\mathbf{y}|\mathbf{u}) f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \right] \\
&= \int_{\mathbb{R}^q} \left[\frac{\partial}{\partial \varphi} f_{\mathbf{Y}|\mathbf{U}}(\mathbf{y}|\mathbf{u}) \right] \frac{f_{\mathbf{U}}(\mathbf{u})}{f_{\mathbf{Y}}(\mathbf{y})} d\mathbf{u} + \int_{\mathbb{R}^q} \frac{f_{\mathbf{Y}|\mathbf{U}}(\mathbf{y}|\mathbf{u})}{f_{\mathbf{Y}}(\mathbf{y})} \left[\frac{\partial}{\partial \varphi} f_{\mathbf{U}}(\mathbf{u}) \right] d\mathbf{u} \\
&= \int_{\mathbb{R}^q} \frac{f_{\mathbf{Y},\mathbf{U}}(\mathbf{y},\mathbf{u})}{f_{\mathbf{Y}}(\mathbf{y})} \frac{1}{f_{\mathbf{U}}(\mathbf{u})} \left[\frac{\partial}{\partial \varphi} f_{\mathbf{U}}(\mathbf{u}) \right] d\mathbf{u} \\
&= \int_{\mathbb{R}^q} f_{\mathbf{U}|\mathbf{Y}}(\mathbf{u}|\mathbf{y}) \left[\frac{\partial}{\partial \varphi} \log f_{\mathbf{U}}(\mathbf{u}) \right] d\mathbf{u} \\
&= E \left[\frac{\partial \log f_{\mathbf{U}}(\mathbf{u})}{\partial \varphi} \middle| \mathbf{y} \right]
\end{aligned}$$

sendo que esta equação não pode ser simplificada sem se especificar uma forma para a distribuição dos efeitos aleatórios, conforme McCulloch e Searle (2001).

4.3.2 Testes de Hipótese

Após feito o ajuste de um modelo para um determinado conjunto de dados, é de interesse do pesquisador verificar a veracidade daquele modelo ajustado, ou seja, realizar testes para os parâmetros do modelo individualmente ou conjuntamente. Isso acontece pelo fato de querer verificar se todos os parâmetros no modelo são significantes ou não, e de querer verificar cada parâmetro individualmente. Os testes mais conhecidos na literatura são os de Wald, escore e o teste da razão de verossimilhança. Todos eles já foram mencionados anteriormente, porém serão introduzidos nesta seção em uma linguagem mais voltada para o modelo proposto nesse capítulo.

4.3.2.1 Teste da razão de verossimilhança

Conforme Pinheiro e Bates (2000), esse método é mais utilizado quando os parâmetros do modelo foram estimados pelo método de máxima verossimilhança. Esse teste é baseado na razão de dois modelos ajustados, e é muito utilizado quando o pesquisador tem interesse

em verificar se um modelo com menos parâmetros está mais bem ajustado do que um com mais parâmetros. Caso os parâmetros tenham sido estimados pelo método de máxima verossimilhança restrita, é importante observar que, o teste só poderá ser aplicado se os dois modelos tiverem sido ajustado pelo mesmo método e também os efeitos fixos tiverem a mesma estrutura.

Sendo L_2 a verossimilhança do modelo completo e L_1 a verossimilhança do modelo restrito, tem-se que $L_2 > L_1$ e por consequência, $\log L_2 > \log L_1$. Conforme McCulloch e Searle (2001), quando se tem interesse em testar se sua única componente da variância é igual a zero, a estatística do teste de razão de verossimilhança será representada por:

$$\xi_{RV} = 2\log(L_2/L_1) = 2(\log L_2 - \log L_1).$$

Na maioria das vezes a verossimilhança não pode ser avaliada analiticamente, o mesmo ocorre com a estatística da razão de verossimilhança.

4.3.2.2 Teste de Wald

Conforme McCulloch e Searle (2001), quando se tem amostras grandes e a construção da informação observada ou esperada é possível, o teste de Wald pode ser feito utilizando a normalidade de grandes amostras dos estimadores. Ele pode ser elaborado para um parâmetro individual, podendo ser representado por:

$$\frac{\hat{\beta}_j - \beta_{j,0}}{\sqrt{\widehat{Var}(\hat{\beta}_i)}} \sim \mathcal{N}(0, 1), \text{ para } n \rightarrow \infty$$

ou para um conjunto de combinações lineares dos parâmetros, podendo ser representado por:

$$\mathbf{K}^T \hat{\boldsymbol{\beta}} - \mathbf{K}^T \hat{\boldsymbol{\beta}}_0 \sim \mathcal{N}(0, \mathbf{K}^T \mathbf{I}^{-1} \mathbf{K}), \text{ para } n \rightarrow \infty$$

sendo \mathbf{I} a matriz de informação observada ou esperada e \mathbf{K} a matriz de posto linha completo.

4.3.2.3 Teste de Escore

Conforme McCulloch e Searle (2001), esse teste verifica a presença de um único efeito aleatório ou vários efeitos aleatórios em um modelo. Esse teste também foi proposto por Commenges *et al.* (1994); Jacquin-Gadda e Commenges (1995); Lin (1997); Commenges

e Jacquin-Gadda (1997). O teste escore tem a vantagem de não exigir os estimadores de máxima verossimilhança em um MLGM. Contudo, esse teste muitas vezes tem menos poder do que os testes baseados nos modelos de efeitos aleatórios.

4.3.3 Verossimilhança perfilada

Este procedimento é muito utilizado quando o modelo a ser ajustado envolve parâmetros de perturbação. A idéia é substituir esse parâmetro por uma estimativa consistente na verossimilhança original, para valores fixados dos parâmetros de interesse, conforme (SILVA, 2005).

A função resultante é chamada de função de verossimilhança perfilada. Ela depende somente dos parâmetros de interesse. Essa função de verossimilhança não pode ser chamada e nem utilizada como uma função de verossimilhança genuína. Portanto, utilizar a função de verossimilhança como sendo uma função genuína pode levar a obter alguns problemas, pois algumas das propriedades da verossimilhança genuína não são válidas para a função de verossimilhança perfilada. Por exemplo, a função escore não tem obrigatoriamente média zero e sua informação pode apresentar vício. Problemas como, por exemplo, inconsistência e ineficiência de estimadores podem ser encontrados, conforme Silva (2005).

Desta forma, muitos estudos foram propostos para ajustar melhor essa verossimilhança perfilada, a fim de reduzir esses problemas citados acima. Conforme (SILVA, 2005), vários pesquisadores propuseram ajustes para a função de verossimilhança perfilada. Para mais detalhes verificar (SILVA, 2005).

4.3.3.1 Função

Considere y_1, \dots, y_n variáveis aleatórias independentes, identicamente distribuídas e pertencente a classe da família exponencial biparamétrica. Sua função densidade ou de probabilidade será dada pela seguinte forma:

$$p(y; \mu, \varphi) = \exp[\alpha_1(\mu, \varphi)d_1(y) + \alpha_2(\mu, \varphi)d_2(y) - \rho(\mu, \varphi) + v(y)], \quad (4.13)$$

sendo que (μ, φ) representa o vetor de parâmetros desconhecidos, $\alpha_1(\cdot, \cdot)$, $\alpha_2(\cdot, \cdot)$, $\rho(\cdot, \cdot)$, $d_1(\cdot)$, $d_2(\cdot)$ e $v(\cdot)$ são funções conhecidas e que o suporte de $p(y; \mu, \varphi)$ não depende de (μ, φ) . Também assume-se que $\alpha_1(\cdot, \cdot)$, $\alpha_2(\cdot, \cdot)$, $\rho(\cdot, \cdot)$ têm as quatro primeiras derivadas contínuas em relação aos componentes do vetor de parâmetros (μ, φ) . Assim, a estatística

suficiente minimal é representada por $\left(\sum_{i=1}^n d_1(y_j), \sum_{i=1}^n d_2(y_j) \right)$. Onde y representa tanto a variável aleatória quanto seu valor observado.

Conforme Silva (2005), sendo $\boldsymbol{\theta}$ o vetor paramétrico de um modelo estatístico hipotético com função de verossimilhança $L(\boldsymbol{\theta})$ e para $\mu = \mu(\boldsymbol{\theta})$ representando o vetor de parâmetros de interesse, a função de verossimilhança perfilada para esse vetor será representada por

$$L_p(\mu) = \sup_{\boldsymbol{\theta}|\mu} L(\boldsymbol{\theta})$$

onde $\boldsymbol{\theta}|\mu$ denota todos os valores de $\boldsymbol{\theta}$ sendo que $\mu(\boldsymbol{\theta}) = \mu$. O logaritmo da função de verossimilhança perfilada é dada por $\ell_p(\mu)$. A função escore perfilada é representada por, $u_p(\mu) = \partial \ell_p(\mu) / \partial \mu$ e o ponto de máximo da função $L_p(\mu)$ é denotada por $\hat{\mu}_p$.

Suponha agora que $\boldsymbol{\theta} = (\mu, \varphi)$, logo a verossimilhança perfilada será representada por:

$$L_p(\mu) = L(\mu, \hat{\varphi}_\mu),$$

onde $\hat{\varphi}_\mu$ é a estimativa de máxima verossimilhança de φ para μ fixado. Ou seja, $\hat{\varphi}_\mu$ é a raiz da equação $\partial \ell(\boldsymbol{\theta}) / \partial \varphi = 0$, sendo $\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$.

Conforme Pace e Salvani (1997), a função de verossimilhança perfilada contém algumas propriedades interessantes. Essas propriedades não são observadas apenas na família exponencial biparamétrica, mais sim de validade geral. Algumas dessas propriedades são dadas por:

- i. $\hat{\mu}_p = \hat{\mu}$, em que $\hat{\mu}$ é a estimativa de máxima verossimilhança de μ e $\hat{\mu}_p$ é a estimativa de máxima verossimilhança perfilada;
- ii. Ao realizar testes de hipóteses sobre μ , a estatística de razão de verossimilhança baseada em $\ell_p(\mu)$ é igual à baseada em $\ell(\mu, \varphi)$, ou seja, a estatística de razão de verossimilhança é dada por:

$$\xi_{RV} = 2[\ell(\hat{\mu}, \hat{\varphi}) - \ell(\mu, \hat{\varphi}_\mu)] = 2[\ell_p(\hat{\mu}) - \ell_p(\mu)],$$

em que $\hat{\varphi}$ é a estimativa de máxima verossimilhança de φ .

Voltando para a classe da família exponencial biparamétrica, onde μ e φ são unidimensionais, o logaritmo da verossimilhança para estes parâmetros utilizando a função de

densidade (4.13) será representada por:

$$\ell = \ell(\mu, \varphi) = \sum_{j=1}^n [\alpha_1(\mu, \varphi)d_1(y_j) + \alpha_2(\mu, \varphi)d_2(y_j) - \rho(\mu, \varphi)].$$

A parcela $\sum_{j=1}^n \varphi(y_j)$ foi desprezada da função de verossimilhança, pois não depende dos parâmetros de interesse.

Conforme Silva (2005), sejam $\alpha_i^{(s,t)}(\mu, \varphi) = \partial^{s+t}\alpha_i(\mu, \varphi)/\partial\mu^s\partial\varphi^t$ para $(i = 1, 2)$ e $\rho^{(s,t)}(\mu, \varphi) = \partial^{(s+t)}\rho(\mu, \varphi)/\partial\mu^s\partial\varphi^t$. Portanto,

$$\frac{\partial\ell}{\partial\mu} = \sum_{j=1}^n \sum_{i=1}^2 [\alpha_i^{(1,0)}(\mu, \varphi)d_i(y_j)] - n\rho^{(1,0)}(\mu, \varphi)$$

e

$$\frac{\partial\ell}{\partial\varphi} = \sum_{j=1}^n \sum_{i=1}^2 [\alpha_i^{(0,1)}(\mu, \varphi)d_i(y_j)] - n\rho^{(0,1)}(\mu, \varphi).$$

Utilizando as definições $E\{\partial\ell/\partial\mu\} = E\{\partial\ell/\partial\varphi\} = 0$ e $\beta_i(\mu, \varphi) = E\{d_i(y)\}$ para $(i = 1, 2)$. Tem-se que:

$$\rho^{(1,0)}(\mu, \varphi) = \sum_{i=1}^2 \alpha_i^{(1,0)}(\mu, \varphi)\beta_i(\mu, \varphi)$$

e

$$\rho^{(0,1)}(\mu, \varphi) = \sum_{i=1}^2 \alpha_i^{(0,1)}(\mu, \varphi)\beta_i(\mu, \varphi).$$

Resolvendo o sistema das duas equações acima, para $\beta_1(\mu, \varphi)$ e $\beta_2(\mu, \varphi)$, encontra-se as seguintes equações

$$\beta_1 = \frac{\rho^{(1,0)}\alpha_2^{(0,1)} - \rho^{(0,1)}\alpha_2^{(1,0)}}{\alpha_1^{(1,0)}\alpha_2^{(0,1)} - \alpha_1^{(0,1)}\alpha_2^{(1,0)}} \quad e \quad \beta_2 = \frac{\rho^{(0,1)}\alpha_2^{(1,0)} - \rho^{(1,0)}\alpha_2^{(0,1)}}{\alpha_1^{(1,0)}\alpha_2^{(0,1)} - \alpha_1^{(0,1)}\alpha_2^{(1,0)}} \quad (4.14)$$

onde assume-se $\alpha_1^{(1,0)}\alpha_2^{(0,1)} - \alpha_1^{(0,1)}\alpha_2^{(1,0)} \neq 0$, afim de que as equações existam.

Conforme Silva (2005), os estimadores de máxima verossimilhança satisfazem simultaneamente:

$$\sum_{j=1}^n \sum_{i=1}^2 [\alpha_i^{(1,0)}(\hat{\mu}, \hat{\varphi})d_i(y_j)] - n\rho^{(1,0)}(\hat{\mu}, \hat{\varphi}) = 0$$

e

$$\sum_{j=1}^n \sum_{i=1}^2 [\alpha_i^{(0,1)}(\hat{\mu}, \hat{\varphi}) d_i(y_j)] - n\rho^{(0,1)}(\hat{\mu}, \hat{\varphi}) = 0.$$

Considerando as duas equações estimadas da verossimilhança e resolvendo o sistema em relação a $\sum_{j=1}^n d_1(y_j)$ e $\sum_{j=1}^n d_2(y_j)$ utilizando a equação (4.14), tem-se que a solução será dada por:

$$\sum_{j=1}^n d_1(y_j) = n\beta_1(\hat{\mu}, \hat{\varphi}) \quad \text{e} \quad \sum_{j=1}^n d_2(y_j) = n\beta_2(\hat{\mu}, \hat{\varphi}).$$

Utilizando as expressões obtidas para $\rho^{(0,1)}(\mu, \varphi)$ e $\rho^{(1,0)}(\mu, \varphi)$, as derivadas $\partial\ell/\partial\varphi$ e $\partial\ell/\partial\mu$ podem ser escrita da seguinte forma:

$$\frac{\partial\ell}{\partial\varphi} = n \sum_{i=1}^2 \alpha_i^{(0,1)}(\mu, \varphi) [\beta_i(\hat{\mu}, \hat{\varphi}) - \beta_i(\mu, \varphi)]$$

e

$$\frac{\partial\ell}{\partial\mu} = n \sum_{i=1}^2 \alpha_i^{(1,0)}(\mu, \varphi) [\beta_i(\hat{\mu}, \hat{\varphi}) - \beta_i(\mu, \varphi)].$$

Assim, considerando μ fixado e depois φ fixado, os estimadores de máxima verossimilhança de $\hat{\varphi}_\mu$ e $\hat{\mu}_\varphi$ satisfazem respectivamente as equações abaixo

$$\sum_{i=1}^2 \alpha_i^{(0,1)}(\mu, \hat{\varphi}_\mu) [\beta_i(\hat{\mu}, \hat{\varphi}) - \beta_i(\mu, \hat{\varphi}_\mu)] = 0 \quad (4.15)$$

$$\sum_{i=1}^2 \alpha_i^{(1,0)}(\hat{\mu}_\varphi, \varphi) [\beta_i(\hat{\mu}, \hat{\varphi}) - \beta_i(\hat{\mu}_\varphi, \varphi)] = 0. \quad (4.16)$$

Portanto, o logaritmo da função de verossimilhança perfilada dos parâmetro μ e φ são dados por

$$\ell_p(\mu) = \ell(\mu, \hat{\varphi}_\mu) = n \left\{ \sum_{i=1}^2 [\alpha_i(\mu, \hat{\varphi}_\mu) \beta_i(\hat{\mu}, \hat{\varphi})] - \rho(\mu, \hat{\varphi}_\mu) \right\}$$

e

$$\ell_p(\varphi) = \ell(\hat{\mu}_\varphi, \varphi) = n \left\{ \sum_{i=1}^2 [\alpha_i(\hat{\mu}_\varphi, \varphi) \beta_i(\hat{\mu}, \hat{\varphi})] - \rho(\hat{\mu}_\varphi, \varphi) \right\}$$

sendo que as estimativas de $\hat{\varphi}_\mu$ e $\hat{\mu}_\varphi$ são as raízes respectivamente das equações (4.15) e (4.16).

5 ANÁLISE DE DADOS DE SUCO DE CAJU

Um experimento foi realizado a fim de comparar a aceitação de cinco marcas comerciais de suco de caju. Nesse experimento 100 (cem) provadores não treinados, previamente selecionados, após provarem uma pequena quantidade de suco foram convidados a responder várias questões, dentre elas: “descreva o quanto você gostou ou desgostou de cada amostra com relação à **impressão global**”. A resposta a esta questão se baseia na seguinte escala ordinal: (1) Gostei muitíssimo; (2) Gostei muito; (3) Gostei moderadamente; (4) Gostei ligeiramente; (5) Nem gostei, nem desgostei; (6) Desgostei ligeiramente; (7) Desgostei moderadamente; (8) Desgostei muito; (9) Desgostei muitíssimo. Para efeito deste trabalho as opções de resposta foram reagrupadas em cinco categorias: (1) Gostei muito (junção das categorias 1 e 2); (2) Gostei moderadamente (junção das categorias 3 e 4); (3) Nem gostei, nem desgostei; (4) Desgostei moderadamente (junção das categorias 6 e 7); (5) Desgostei muito (junção das categorias 8 e 9).

O suco provado era proveniente de amostras de suco tropical de caju pronto para beber, armazenadas em embalagens Tetrabrik de 1 litro. Foram selecionadas 5 (cinco) embalagens de cada marca pertencentes a um mesmo lote, coletadas em um supermercado de Fortaleza - CE. O experimento foi realizado entre o segundo semestre do ano de 2008 e o primeiro semestre do ano de 2009.

O experimento foi realizado de acordo com Macfie *et al.* (1989), de forma que todas as marcas de caju foram provadas o mesmo número de vezes e foram precedidas pelas demais também a mesma quantidade de vezes e ocuparam todas as posições na ordem de prova. Os provadores foram questionados quanto ao sexo, faixa etária, formada pelas categorias menos de 18 anos, de 19 a 25 anos, de 26 a 35 anos, de 36 a 45 anos, de 46 a 55 anos e acima de 56 anos; quanto ao consumo de suco de caju e quanto ao consumo por outros sucos de frutas, cujas opções eram (1) Consumo muito frequentemente, (2) Consumo frequentemente, (3) Consumo ocasionalmente, (4) Consumo pouco e (5) Não consumo e também quanto ao gosto por suco de caju e por outros sucos de frutas cujas respostas para estas duas últimas perguntas possíveis eram (1) Gosto muitíssimo, (2) Gosto muito e (3)

Gosto moderadamente. Essas variáveis podem exercer influência sobre a impressão global com relação ao suco de caju e foram aqui consideradas como covariáveis. O questionário dessa pesquisa pode ser visto no Anexo.

Os dados foram analisados utilizando-se o modelo de chances proporcionais mistos. Inicialmente foi feita a análise do modelo sem o efeito aleatório, ou seja, ajustou-se modelo de chances proporcionais com efeitos fixos e posteriormente comparado com o modelo com a adição do efeito aleatório. Os resultados da análise descritiva e do ajuste dos modelos de efeitos fixos e aleatórios serão apresentados a partir da seção 5.2.

Atualmente, os métodos computacionais são essenciais nas diversas áreas de atuação, sendo assim, não seria diferente a grande importância destes na área da estatística. Ao resolver problemas nessa área, encontram-se equações muito extensas e complicadas, portanto não sendo viável tentar resolver esse problema braçalmente, ou, muitas vezes não é possível. Contudo, cresce o número de *softwares* estatísticos para tentar solucionar esses problemas. São exemplos de *softwares*: SAS, StatPlus, R Project, IBM SPSS, Stata, Minitab, Reviews, S-PLUS, Multilog, entre outros.

O pacote ordinal utilizado nesse estudo está implementado no *software* R project, que por sua vez está disponível livremente pela *Free Software Foundation's GNU General Public License* para uso diário. O R é uma linguagem e também um ambiente para fazer cálculos estatísticos e gráficos. Ele ainda oferece uma ampla variedade de métodos estatísticos como: análise descritiva, análise gráfica, modelagem linear e não linear, testes paramétricos e não-paramétricos, análise de séries temporais, entre outros. Para mais informações acessar o site <http://www.r-project.org/>.

5.1 Pacote Ordinal

O pacote ordinal teve sua primeira versão publicada em Março de 2010 e após essa versão muitas mudanças foram propostas. Nesses últimos anos ele obteve várias atualizações que levaram a grandes avanços nas análises dos modelos. Em fevereiro de 2012, Rune Haubo publicou a última versão para esse pacote que está disponível a partir da versão 2.13 do R, conforme Christensen (2012). Esse pacote tem como característica implementar modelos de ligação cumulativos e modelos de ligação cumulativos mistos, logo o modelo de chances proporcionais se adequa nesse caso.

O pacote traz duas funções para ajustes de modelos, a função *cumulative link models* (*clm*) e *cumulative link mixed models* (*clmm2*).

A função `clm` é usada quando se tem o interesse de ajustar um modelo com apenas efeitos fixos. A estimação desse modelo na função é feita através da função de verossimilhança com a utilização do algoritmo de *Newton Raphson* para otimizar a função. Já na função `clmm2`, é de interesse do pesquisador ajustar um modelo como efeitos aleatórios. A estimação desse modelo na função é feita através da aproximação de Laplace ou métodos de quadratura de Gauss-Hermite adaptada.

5.2 Análise descritiva

Com o intuito de verificar o comportamento das covariáveis consumo de suco de caju; consumo de outras frutas; gosto de suco de caju e gosto de outras frutas sobre a variável dependente impressão global, foram feitos cruzamentos da variável dependente com cada covariável fixando a variável marca. A representação gráfica desses cruzamentos pode ser vista nas Figuras 3, 4, 5 e 6.

Com relação as Figuras 3 e 4 pode-se observar que a marca A teve um maior número de pessoas que a avaliaram positivamente em todos os grupos das covariáveis “Consumo por suco de caju”, ou “Consumo por outros sucos de frutas”. Já para a marca C aconteceu o contrário, tendo menor aceitação. Ela obteve maiores para “Desgostei muito”, nas categorias “Consumo pouco” e “Consumo ocasionalmente”, isso pra covariável consumo de suco de caju. Pra consumo de outros sucos “Desgostei muito” é mais frequente em “Muito frequentemente”.

Com relação as Figuras 5 e 6 pode-se observar que a marca A teve um maior número de pessoas que avaliaram positivamente em todos os grupos das covariáveis “Gosto por suco de caju”, ou “Gosto por outros sucos de frutas”. Já para a marca C aconteceu o contrário, tendo menor aceitação. Ela teve maiores frequência nas categorias “Desgostei muito” e “Desgostei moderadamente”, isso para as covariáveis gosto por suco de caju e por outros sucos de frutas.

Para mais detalhes sobre os valores exatos das frequências referentes à Figuras 3, 4, 5, e 6, ver as Tabelas 5, 6, 7 e 8 encontradas no Apêndice.

5.3 Modelo de efeitos fixos

Ajustando inicialmente o modelo apresentado em (3.15), em que o vetor de covariáveis $\mathbf{x} = (x_1, x_2, \dots, x_7)$, sendo:

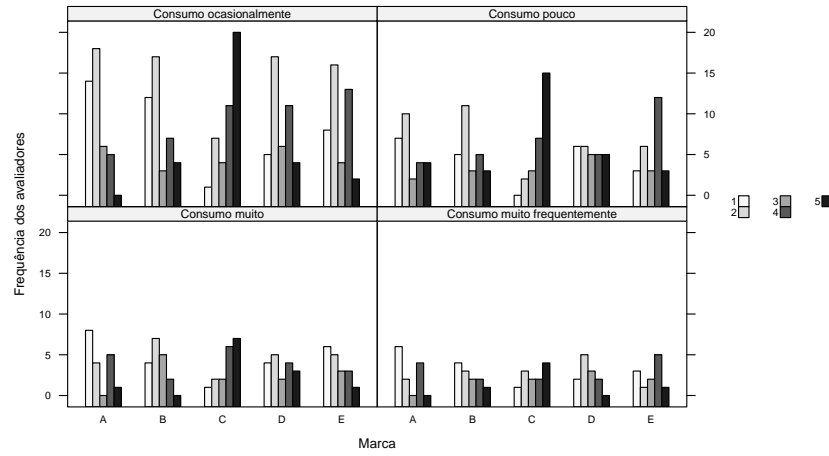


Figura 3: Cruzamento das variáveis marca, consumo por suco de caju e impressão global.

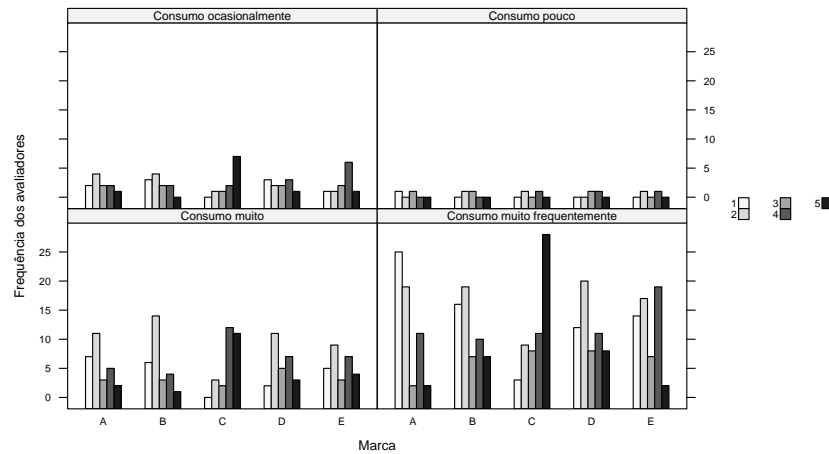


Figura 4: Cruzamento das variáveis marca, consumo por outros sucos de frutas e impressão global.

- x_1 representa idade;

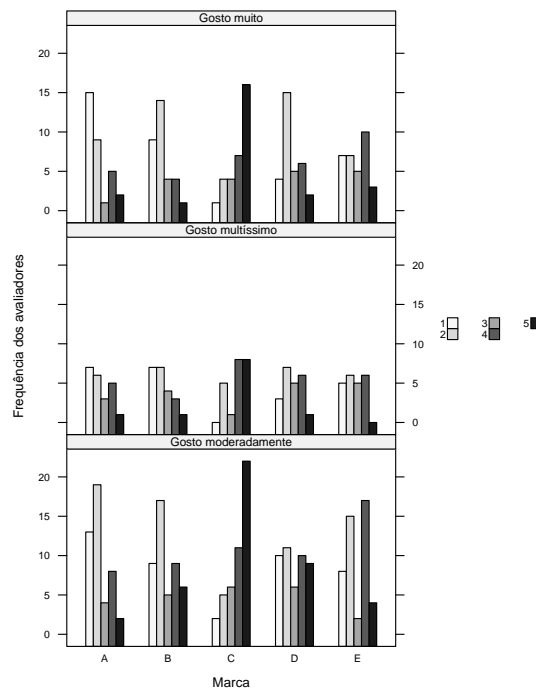


Figura 5: Cruzamento das variáveis marca, gosto por suco de caju e impressão global.

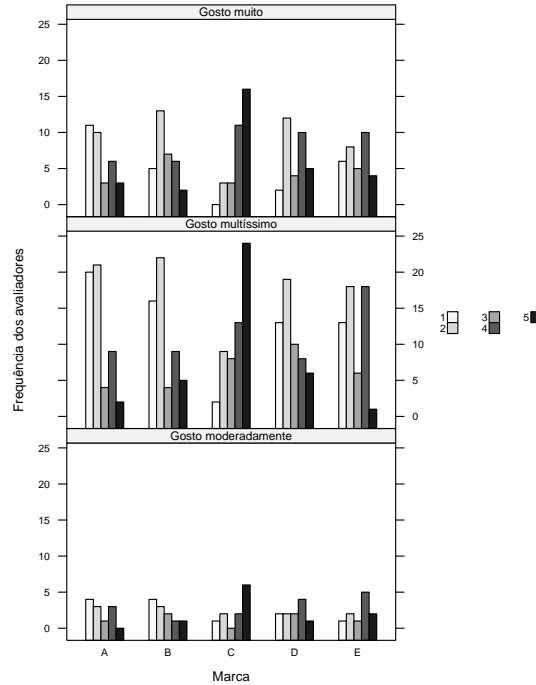


Figura 6: Cruzamento das variáveis marca, gosto por outros sucos de frutas e impressão global

- x_2 , sexo; x_3 , consumo de suco de caju;

- x_4 , consumo de outros sucos de frutas;
- x_5 , gosto por suco de caju;
- x_6 , gosto por outros sucos de frutas;
- x_7 , marca do suco de caju provado;
- Y representa a nota dada a impressão global do suco de caju;
- $\gamma_h(\mathbf{x}) = P(Y \leq h)$, onde $h = 1, 2, \dots, 5$

no qual foram consideradas como covariáveis idade, sexo, consumo por suco de caju, consumo por outras frutas, gosto por suco de caju, gosto por outras frutas e marca. Apenas as covariáveis marca e gosto por outros sucos de frutas (que foi aqui denotada por “Outros”) foram significantes. Os resultados mostraram que o AIC e BIC do modelo com apenas as covariáveis marca e gosto de outros sucos de frutas ($AIC = 1469,88$ e $BIC = 1512,02$) foi menor do que a do modelo com todas as covariáveis ($AIC = 1477,17$ e $BIC = 1561,42$). Os valores estimados para o modelo com apenas duas covariáveis podem ser visto na Tabela 3.

Tabela 3: Estimativas para os parâmetros do modelo de chances proporcionais para impressão global do suco de caju.

Parâmetro	Estimativa	Erro-padrão	Valor z	Nível descritivo
Interceptos				
β_{02}	-0,485	0,203	-2,388	
β_{03}	1,046	0,208	5,021	
β_{04}	1,617	0,215	7,516	
β_{05}	3,084	0,246	12,503	
Covariáveis				
marca(B)	0,324	0,257	1,264	0,206
marca(C)	2,585	0,281	9,176	< 0,001
marca(D)	0,819	0,257	3,180	0,001
marca(E)	0,824	0,258	3,192	0,001
outros (2)	0,498	0,177	2,806	0,005
outros (3)	0,449	0,273	1,645	0,099

A Figura 7 representa as probabilidades estimadas de cada categoria de resposta para diferentes combinações de marcas de suco e opções de resposta para “outros”. Por ela, observa-se que a marca de suco C obteve a pior avaliação pois apresenta as maiores probabilidades associadas à categoria 5, que corresponde a desgostar muito da impressão global, o que não ocorre com as demais marcas. Ainda observou-se que a marca A foi a mais

aceita em relação às outras marcas, pois apresentou maiores probabilidades associadas às categorias 1 e 2, que correspondem a gostar muito e moderadamente da impressão global, respectivamente. Pode-se observar ainda que em todas as marcas obteve-se uma maior probabilidade de serem melhor avaliadas na categoria “1” de outros, ou seja, quem gosta muitíssimo de outros sucos de frutas.

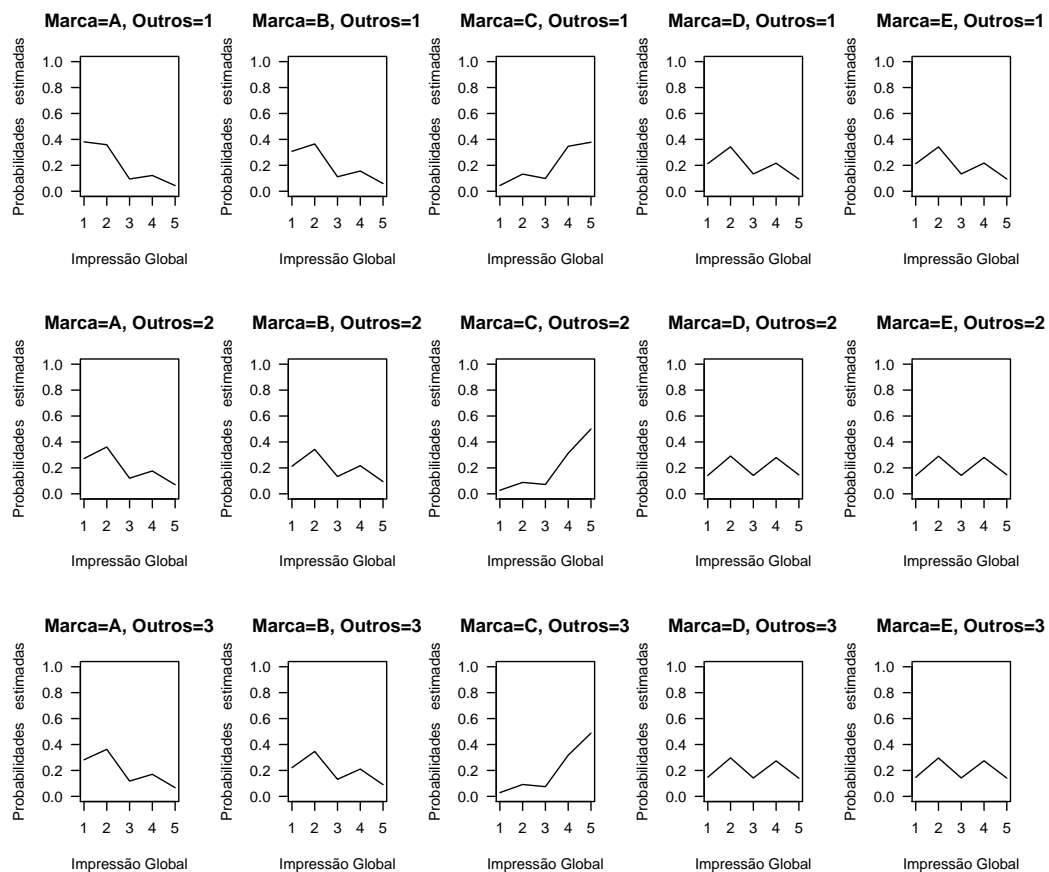


Figura 7: Estimativas das probabilidades referentes ao modelo ajustado sem o efeito aleatório.

5.4 Modelo de efeitos aleatórios

Foi ajustado um modelo de chances proporcionais misto (4.7) no qual u_i representa o efeito do avaliador i . As covariáveis consideradas foram as mesmas do modelo anterior. Com base nesse ajuste pode-se observar que as variáveis idade, sexo, consumo de suco de caju, consumo de outros sucos de frutas e gosto de suco de caju não foram significantes para o modelo. Assim, o modelo resultante contém apenas duas covariáveis sendo estas marca e gosto por outros sucos de frutas “outros” que também foram significativas no modelo de efeitos fixos. Os resultados mostraram ainda que o AIC e BIC do modelo

com apenas as covariáveis marca e gosto por outros sucos de frutas ($AIC = 1436,15$ e $BIC = 1482,51$) foi menor do que a do modelo com todas as covariáveis ($AIC = 1448,12$ e $BIC = 1536,62$).

Ao ajustar modelo com as duas variáveis significativas e o efeito aleatório pode-se observar que a estimativa para o desvio padrão do coeficiente aleatório foi 0,997. Os intervalos de 95% e 99% de confiança para σ_u não contém o valor zero, indicando a existência de efeito aleatório, o que pode ser observado pela Figura 8 no qual se encontra a representação gráfica do comportamento da verossimilhança perfilada. Foi verificado a existência de efeito aleatório no modelo completo e neste o efeito também era significativo.

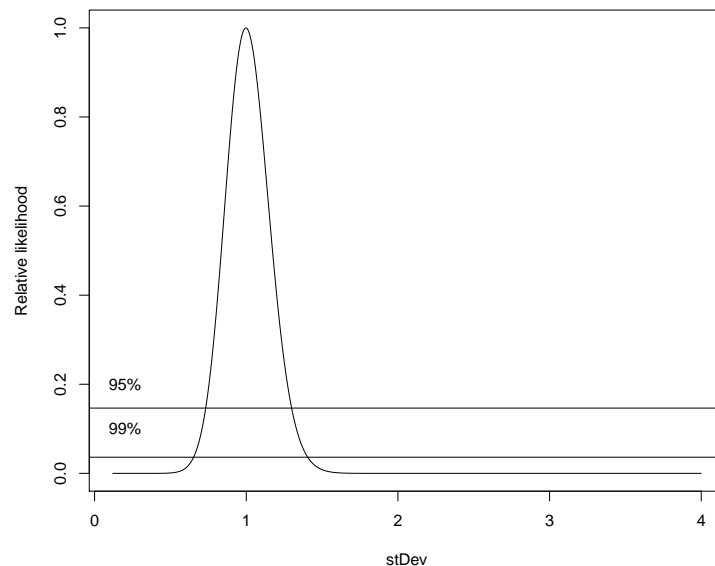


Figura 8: Verossimilhança perfilada de σ_u do modelo (4.1) para os dados de suco de caju.

Os valores estimados para esse modelo estão apresentados na Tabela 4.

A Figura 9 apresenta as probabilidades estimadas de cada categoria de resposta para diferentes combinações de marcas de suco e opções de resposta para “outros”. Por ela observa-se que a marca de suco C obteve a pior avaliação pois apresenta as maiores probabilidades associadas à categoria 5, que corresponde a desgostar muito da impressão global, o que não ocorre com as demais marcas. Ainda se observou que a marca A foi a mais aceita em relação as outras marcas, pois apresentou maiores probabilidades associadas as categorias 1 e 2, que correspondem a gostar muito e moderadamente da impressão global respectivamente. Pode-se observar ainda que em todas as marcas teve-

Tabela 4: Estimativas para os parâmetros do modelo de chances proporcionais para impressão global do suco de caju.

Parâmetro	Estimativa	Erro-padrão	Valor z	Nível descritivo
Interceptos				
β_{02}	-0,571	0,252	-2,266	
β_{03}	1,230	0,258	4,755	
β_{04}	1,904	0,267	7,110	
β_{05}	3,615	0,309	11,690	
Covariáveis				
marca(B)	0,384	0,267	1,439	0,150
marca(C)	3,000	0,303	9,887	< 0,001
marca(D)	0,999	0,268	3,720	< 0,001
marca(E)	0,958	0,268	3,570	< 0,001
outros (2)	0,626	0,287	2,181	0,029
outros (3)	0,445	0,436	1,021	0,307

se uma maior probabilidade de serem melhor avaliadas na categoria “1” de “outros”, ou seja, quem gosta muitíssimo de outros sucos de frutas.

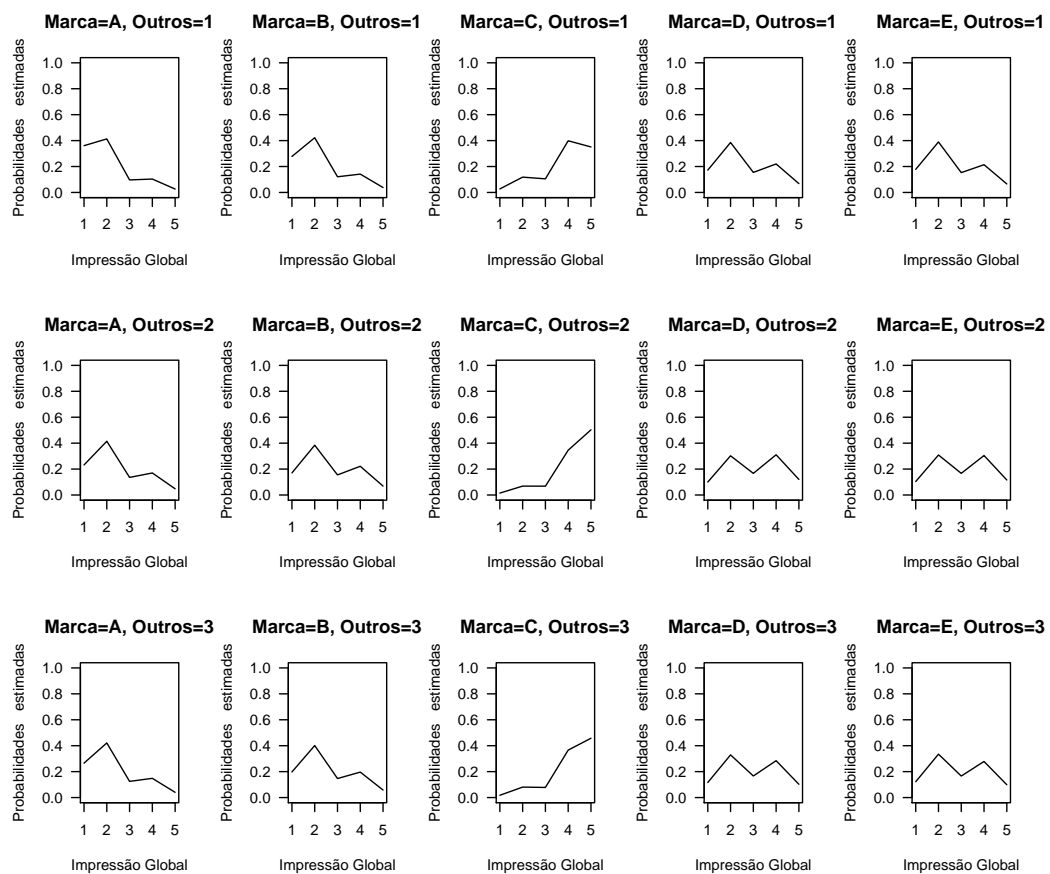


Figura 9: Estimativas das probabilidades referentes ao modelo ajustado sem o efeito aleatório.

6 CONSIDERAÇÕES FINAIS

O principal objetivo desse estudo foi modelar a possível dependência da variável impressão global com as demais covariáveis. Ao analisar o banco de dados constatou-se que havia observações repetidas para uma mesma unidade experimental (avaliador). Assim, ajustar um modelo com efeitos aleatórios foi necessário para verificar melhor esse comportamento.

Ao realizar a análise descritiva com marca e impressão global fixadas, pode-se observar que a marca A teve um maior número de pessoas que a avaliaram positivamente em todos os grupos das covariáveis “Consumo por suco de caju”, “Consumo por outros sucos de frutas”, “Gosto por suco de caju” e “Gosto por outros sucos de frutas”. Já para a marca C aconteceu o contrário, tendo menor aceitação. Ela obteve maiores frequências para as categorias “Desgostei muito” e “Desgostei moderadamente” nas respectivas covariáveis “Consumo por suco de caju”, “Consumo por outros sucos de frutas”, “Gosto por suco de caju” e “Gosto por outros sucos de frutas”.

Ao ajustar um modelo com as metodologias apresentadas neste estudo, verificou-se que aquele que continha adição de um efeito aleatório possibilitava a modelagem da dependência entre as observações repetidas de um mesmo provador, e que este era significativo no modelo. Os dois modelos ajustados retornaram valores próximos, porém o erro padrão dos estimadores dos parâmetros é maior no modelo de efeitos aleatórios, o que indica que o modelo de efeitos fixos pode estar subestimando a variância dos estimadores. Com relação aos dados obtidos do modelo com efeito aleatório, as probabilidades estimadas indicaram uma baixa aceitação à marca C e alta aceitação da marca A. Já para as demais marcas, houve uma divisão na opinião dos provadores.

O uso do *software* R 2.13.1 foi de suma importância, dada a complexidade dos algoritmos utilizados, facilitando o desenvolvimento desse trabalho.

Como sugestão para trabalhos futuros deixa-se a abordagem da análise de diagnóstico em modelos de chance proporcionais com efeito aleatório.

REFERÊNCIAS

1. AGRESTI, A. **Categorical data Analysis**. 2nd ed. Hoboken, N.J.: Wiley-Interscience, c2002. xv, 710 p. (Wiley series in probability and statistics) ISBN 9780471360933 (enc.), 2002.
2. AIUBE, F. **Econometria para Séries Financeiras**. Rio de Janeiro, PUC-RIO, 2007, 51 p.
3. AKAIKE, H. **A new look at statistical model identification**. IEEE Transaction on Automatic Control AC 19, 1974,716-23p.
4. CHRISTENSEN, R. **Analysis of sensory ratings data with Cumulative Link Models**. Journal de la Societé Française de Statistique Dinamarca, 2012. (Submetido)
5. COMMENGES, D.; JACQCMIN-GADDA, H. **Generalized score test of homogeneity based on correlated random effects models**. Journal of the Royal Statistical Society, Series B, v. 59, 1997, p.157-171.
6. COMMENGES, D.; LETENNEUR, L.; JACQMIN, H.; MOREAU, T.; DARTIGUES, J. **Test of homogeneity of binary data with explanatory variables**. Biometrics, v.50, 1994, p.613-620.
7. COSTA, T. **Modelos lineares mistos: uma aplicação na produção de leite de vacas da raça sindi**. Recife, UFRPE, (Dissertação), 2010, 79 p.
8. CUNHA, L. **Modelos Rasch e Escalas de Likert e Thurstone na medição de atitudes**. Lisboa, UL, (Dissertação), 2007, 78p.
9. DAVIS, C. **Statistical methods for the analysis of repeated measurements**. New York: Springer-Verlag, 2002, 442p.
10. DEMÉTRIO, C. **Modelos Lineares Generalizados em Experimentação Agronômica**. Piracicaba, ESALQ/USP, 2002, 113 p.
11. DEMIDENKO, E. **Mixed Models: Theory and Applications**. Hoboken, N.J.: Wiley & Sons, Inc., ISBN 0471601616, 2004, 704 p.

12. DOBSON, A.; BARNETT, A. **An Introduction to Generalized Linear Models**. 3rd ed. Boca Raton, Fl: Chapman & Hall/CRC, 2008. 307 p. : ISBN 9781584889502 (broch.).
13. FIGUEIRA, C. **Modelos Regressão Logística**. Porto Alegre, UFRGS, (Dissertação), 2006, 138p.
14. HARVILLE, D. **Extension of the Gauss-Markov Theorem to include the estimation of random effects**. The Annals of Statistics, 1976, 1598-1608 p.
15. HILDEN-MILTON, J. **Multilevel Diagnostics for Mixed and Hierarchical Linear Models**. University of California, Los Angeles, (Tese), 1995, 120 p.
16. HOFFMANN, R. **Análise de regressão: uma introdução à econometria**. 4. ed. São Paulo: Hucitec, 2006. 378 p. : (Obras didaticas. Colecao economia e planejamento) ISBN 8527100231 (broch.)
17. JØRGENSEN, B.; LABOURIAU, R. **Exponential Families e Theoretical Inference**. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1992, 264p.
18. JACQMIN-GADDA, H.; COMMENGES, D. **Tests of homogeneity for generalized linear models**. Journal of the American Statistical Association, v.90, 1995, p.1237-1246.
19. LAIRD, N.; WARE, J. **Random-effects models for longitudinal data**. Biometrics, 1982, 38, 963-974p.
20. LARA, I. **Modelos de Transição para dados Binários**. Piracicaba, ESALQ/USP, (Tese), 2007, 128p.
21. LEE, Y.; NELDER, J. **Hierarchical generalized linear models**. *Journal of the Royal Statistical Society. Serie B (Methodological)* , 58, 1996, 619-678p.
22. LEMOS, T. **Avaliação de blends de hidrocoloides na estabilização do néctar de caju**, UFC, (Tese), 2010, 11p., RENORBIO, Fortaleza, 2010.
23. LIN, X. **Variance component testing in generalised linear models with random effects**. Biometrika, v.84, 1997, p.309-326.
24. MACFIE, H.; BRATCHELL, N.; GREENHOFF, K.; VALLIS, L. **Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests**. Journal of Sensory Studies v. 4, n. 2, 1989, p. 129-148.

25. MARÔCO, J. **Análise Estatística com o SPSS Statistics**. 5 ed. ReportNumber, 2011, 990p.
26. McCULLAGH, P. **Regression Models for Ordinal data (with discussion)**. *Journal of the Royal Statistical Society. Serie B (Methodological)*, 42, 1980, 109-142p.
27. McCULLOCH, C.; SEARLE, S. **Generalized, Linear, and Mixed Models**. New York: John Wiley & Sons, Inc., 2001, 346 p.
28. NELDER, J.; WEDDERBURN, R. **Generalized linear models**. *Journal of the Royal Statistical Society A*, 135, 3, 1972, 370-84p.
29. NOBRE, J. **Métodos de diagnóstico para modelos lineares mistos**. São Paulo, IME/USP, (Dissertação), 2004, 99p.
30. OKURA, R. **Modelos Regressão para Variáveis Categóricas Ordinais com Aplicações ao problema de classificação**. São Paulo, IME/USP, (Dissertação), 2008, 104p.
31. PACE, L.; SALVAN, A. **Principles of Statistical Inference from a Neo-Fisherian Perspective**. Singapore: World Scientific, 1997.
32. PAULA, G. **Modelos de Regressão com apoio Computacional**. São Paulo: IME, 2010, 403 p.
33. PINHEIRO, J.; BATES, D. **Mixed-effects models in S and S-PLUS**. New York: Springer-Verlag, 2000, 528 p.
34. SCHWARZ, G. **Estimation the dimension of a model**, *Annals of Statistics* 6, 1978, 461-4p.
35. SILVA, M. **Estimação e teste de hipótese baseados em verossimilhanças perfiladas** São Paulo, IME/USP,(Tese), 2005, 247 p.
36. STRAM, D.; WEI, L.; WARE, J. **Analysis of repeated ordered categorical outcomes with possibly missing observations and time-dependent covariates**. *Journal of American Statistical Association*, Boston, v. 83, p. 631 – 637, 1988.
37. TURKMAN, M.; SILVA, G. **Modelos Lineares Generalizados - da teoria à prática**. Lisboa, UL/UTL, 2000, 151 p.

APÊNDICE

Tabela 5: Cruzamento das variáveis impressão global, marca e consumo de suco de caju.

Consumo de suco de caju	Marca	Impressão Global										Total
		1		2		3		4		5		
Consumo muito frequentemente	A	6	6%	2	2%	0	0%	4	4%	0	0%	12
	B	4	4%	3	3%	2	2%	2	2%	1	1%	
	C	1	1%	3	3%	2	2%	2	2%	4	4%	
	D	2	2%	5	5%	3	3%	2	2%	0	0%	
	E	3	3%	1	1%	5	5%	5	5%	1	1%	
Consumo muito	A	8	8%	4	4%	0	0%	5	5%	1	1%	18
	B	4	4%	7	7%	5	5%	2	2%	0	0%	
	C	1	1%	2	2%	2	2%	6	6%	7	7%	
	D	4	4%	5	5%	2	2%	4	4%	3	3%	
	E	6	6%	5	5%	3	3%	3	3%	1	1%	
Consumo ocasionalmente	A	14	14%	18	18%	6	6%	5	5%	0	0%	43
	B	12	12%	17	17%	3	3%	7	7%	4	4%	
	C	1	1%	7	7%	4	4%	11	11%	20	20%	
	D	5	5%	17	17%	6	6%	11	11%	4	4%	
	E	8	8%	16	16%	4	4%	13	13%	2	2%	
Consumo pouco	A	7	7%	10	10%	2	2%	4	4%	4	4%	23
	B	5	5%	11	11%	3	3%	5	5%	3	3%	
	C	0	0%	2	2%	3	3%	7	7%	15	15%	
	D	6	6%	6	6%	5	5%	5	5%	5	5%	
	E	3	3%	6	6%	3	3%	12	12%	3	3%	

Tabela 6: Cruzamento das variáveis impressão global, marca e consumo por outros sucos de frutas.

Consumo por outros sucos de frutas	Marca	Impressão Global										Total
		1		2		3		4		5		
Consumo muito frequentemente	A	25	25%	19	19%	2	2%	11	11%	2	2%	59
	B	16	16%	19	19%	7	7%	10	10%	7	7%	
	C	3	3%	9	9%	8	8%	11	11%	28	28%	
	D	12	12%	20	20%	8	8%	11	11%	8	8%	
	E	14	14%	17	17%	7	7%	19	19%	2	2%	
Consumo muito	A	7	7%	11	11%	3	3%	5	5%	2	2%	28
	B	6	6%	14	14%	3	3%	4	4%	1	1%	
	C	0	0%	3	3%	2	2%	12	12%	11	11%	
	D	2	2%	11	11%	5	5%	7	7%	3	3%	
	E	5	5%	9	9%	3	3%	7	7%	4	4%	
Consumo ocasionalmente	A	2	2%	4	4%	2	2%	2	2%	1	1%	11
	B	3	3%	4	4%	2	2%	2	2%	0	0%	
	C	0	0%	1	1%	1	1%	2	2%	7	7%	
	D	3	3%	2	2%	2	2%	3	3%	1	1%	
	E	1	1%	1	1%	2	2%	6	6%	1	1%	
Consumo pouco	A	1	1%	0	0%	1	1%	0	0%	0	0%	2
	B	0	0%	1	1%	1	1%	0	0%	0	0%	
	C	0	0%	1	1%	0	0%	1	1%	0	0%	
	D	0	0%	0	0%	1	1%	1	1%	0	0%	
	E	0	0%	1	1%	0	0%	1	1%	0	0%	

Tabela 7: Cruzamento das variáveis impressão global, marca e gosto de suco de caju.

Gosto de suco de caju	Marca	Impressão Global										Total
		1		2		3		4		5		
Gosto muitíssimo	A	7	7%	6	6%	3	3%	5	5%	1	1%	22
	B	7	7%	7	7%	4	4%	3	3%	1	1%	
	C	0	0%	5	5%	1	1%	8	8%	8	8%	
	D	3	3%	7	7%	5	5%	6	6%	1	1%	
	E	5	5%	6	6%	5	5%	6	6%	0	0%	
Gosto muito	A	15	15%	9	9%	1	1%	5	5%	2	2%	32
	B	9	9%	14	14%	4	4%	4	4%	1	1%	
	C	1	1%	4	4%	4	4%	7	7%	16	16%	
	D	4	4%	15	15%	5	5%	6	6%	2	2%	
	E	7	7%	7	7%	5	5%	10	10%	3	3%	
Gosto moderadamente	A	13	13%	19	19%	4	4%	8	8%	2	2%	46
	B	9	9%	17	17%	5	5%	9	9%	6	6%	
	C	2	2%	5	5%	6	6%	11	11%	22	22%	
	D	10	10%	11	11%	6	6%	10	10%	9	9%	
	E	8	8%	15	15%	2	2%	17	17%	4	4%	

Tabela 8: Cruzamento das variáveis impressão global, marca e gosto por outros sucos de frutas.

Gosto por outros sucos de frutas	Marca	Impressão Global										Total
		1		2		3		4		5		
Gosto muitíssimo	A	20	20%	21	21%	4	4%	9	9%	2	2%	56
	B	16	16%	22	22%	4	4%	9	9%	5	5%	
	C	2	2%	9	9%	8	8%	13	13%	24	24%	
	D	13	13%	19	19%	10	10%	8	8%	6	6%	
	E	13	13%	18	18%	6	6%	18	18%	1	1%	
Gosto muito	A	11	11%	10	10%	3	3%	6	6%	3	3%	33
	B	5	5%	13	13%	7	7%	6	6%	2	2%	
	C	0	0%	3	3%	3	3%	11	11%	16	16%	
	D	2	2%	12	2%	4	4%	10	10%	5	5%	
	E	6	6%	8	8%	5	5%	10	10%	4	4%	
Gosto moderadamente	A	4	4%	3	3%	1	1%	3	3%	0	0%	11
	B	4	4%	3	3%	2	2%	1	1%	1	1%	
	C	1	1%	3	3%	0	0%	2	2%	6	6%	
	D	2	2%	2	2%	2	2%	4	4%	1	1%	
	E	1	1%	2	2%	1	1%	5	5%	2	2%	

Comandos R

```

dados = read.table("suco1.txt", h = TRUE); dados
names(dados)
attach(dados)
require(ordinal)
#####
#####SUCO#####
#####
#"id"                "marca"
#"sexo"              "idade"
#"consumodesuodecaju"  "consumodeoutrossucosdefrutas"
#"gostodesuodecaju"   "gostodeoutrossuodefrutas"
#"cor"                "homogeneidade"
#"aroma"              "sabor"
#"corpo"              "impressãoglobal"
#"atitudedeconsumo"
id=as.factor(dados$id)
sexo=as.factor(dados$sexo)
consumodesuodecaju=as.factor(dados$consumodesuodecaju)
gostodesuodecaju=as.factor(dados$gostodesuodecaju)
cor=as.factor(dados$cor)
aroma=as.factor(dados$aroma)
corpo=as.factor(dados$corpo)
atitudedeconsumo=as.factor(dados$atitudedeconsumo)
consumodeoutrossucosdefrutas=as.factor(dados$consumodeoutros
sucosdefrutas)
gostodeoutrossuodefrutas=as.factor(dados$gostodeoutrossuodefrutas)
homogeneidade=as.factor(dados$homogeneidade)
sabor=as.factor(dados$sabor)
impressãoglobal=as.factor(dados$impressãoglobal)
#####
###SEM EFEITO###
#####
#####
###MODELO COMPLETO##

```

```
#####
#####
#IMPRESSÃO GLOBAL##
#####
mod1=clm(impressãoglobal~marca+idade+sexo+consumodesucodecaju+
gostodesucodecaju+consumodeoutrossucosdefrutas+
gostodeoutrossucodefrutas)
mod1
summary(mod1)
#####
#TIRANDO IDADE, SEXO#####
#GOSTODESUCODECAJU#####
#CONSUMODEOUTROSSUCOSDEFRUTAS###
#####
mod2=clm(impressãoglobal~marca+consumodesucodecaju+
gostodeoutrossucodefrutas)
mod2
summary(mod2)
#####
#TIRANDO CONSUMODESUCODECAJU####
#####
mod3=clm(impressãoglobal~marca+gostodeoutrossucodefrutas)
mod3
summary(mod3)
#####
#SELEÇÃO DE VARIÁVEIS#
#####
AIC(mod1)
AIC(mod2)
AIC(mod3)
BIC(mod1)
BIC(mod2)
BIC(mod3)
#####
#ESTIMATIVAS DE PROBABILIDADES#
#####
```

```

predict(mod3, type="class")
novosDados <- expand.grid(marca=levels(marca),
gostodeoutrossuocodefrutas=levels (gostodeoutrossuocodefrutas))
cbind(novosDados, predict(mod3, newdata = novosDados)$fit)
head(do.call("cbind", predict(mod3, se.fit = TRUE, interval = TRUE)))
#####
# GRÁFICO DAS PROBABILIDADES #
#####
#### ANTES DEFINIR pred.mat####
#####
#x11=c (0.38102378,0.30787732,0.04431751,
0.21339480,0.21221465,0.27219271,0.21275760,
0 .02740168,0.14149800,0.14064438,0.28202340,
0.22109378,0.02874046,0.14756548,0.1 4668151)
#x22=c (0.35897516,0.36497326,0.13223554,
0.34301049,0.34245115,0.36139567,0.34270952,
0 .08784860,0.29098314,0.29010849,0.36288241,
0.34645678,0.09160957,0.29699929,0.2 9614438)
#x33=c (0.09446633,0.11176227,0.09868025,
0.13319102,0.13342052,0.12026846,0.13331515,
0 .07221703,0.14193981,0.14194498,0.11794413,
0.13165369,0.07470808,0.14180806,0.1 4183724)
#x44=c (0.1217727,0.1558283,0.3468501,0.2163287,
0.2172365,0.1760931,0.2168183,0.3125526,0.2796063,
0.2804489,0.1702305,0.2105093,0.3172291,0.2736671,
0.2745267)#x55=c (0.04376199,0.05955883,0.37791663,
0.09407503,0.09467715,0.07005008,0.09439940,
0 .49998010,0.14597277,0.14685321,0.06691960,0.09028643,
0.48771282,0.13996003,0.1 4081018)

pred.mat1=(predict(mod3, newdata = novosDados)$fit)
summary(pred.mat1)
lab <- paste("Gosto=", rep(levels(gostodeoutrossuocodefrutas), 5), ", ",
"marca=", rep(levels(marca), each = 3), sep = "")
lab <- paste("Marca=", rep(levels(marca), 3), ", ",
"Outros=", rep(levels(gostodeoutrossuocodefrutas), each = 5), sep = "")

```

```

par(mfrow = c(3, 5))
for (k in 1:15) {
plot(1:5, pred.mat1[k,], lty = 1, type = "l", ylim = c(0,
1), xlab = "Impressão Global", axis = FALSE, ylab = "Probabilidades
estimadas",foont.ylab=7,
main = lab[ceiling(k)], las = 1)
#axis(1)
#axis(2)
}

#####
###COM EFEITO###
#####
mod5=clmm2(impressãoglobal~marca+idade+sexo+consumodesuodecaju
+gostodesuodecaju
+consumodeoutrossucosdefrutas+gostodeoutrossuodefrutas,
random=id, Hess=TRUE,nAGQ=15)
mod5
summary(mod5)

#Idade, sexo, consumodeoutrossucosdefrutas e gostodesuodecaju
mod6=clmm2(impressãoglobal~marca+consumodesuodecaju+
gostodeoutrossuodefrutas,random=id,
Hess=TRUE,nAGQ=15)
mod6
summary(mod6)

# consumodesuodecaju
mod4=clmm2(impressãoglobal~marca+gostodeoutrossuodefrutas,
random=id, Hess=TRUE,nAGQ=15)
mod4
summary(mod4)

#Variância
pr2 <- profile(mod4, range = c(0.12, 4), nSteps = 20, trace = 0)

```

```

confint(pr2)
plot(pr2)
text(0.2,0.2,"95%")
text(0.2,0.1,"99%")

#####
####CALCULO DAS PROBABILIDADES#
#####
marg=function(marca,gosto,categoria){
i=categoria
#(marca==1 & gosto==1 & categoria==i)
if(marca==1 & gosto==1 ){
if(i==1){
probacum=exp(mod4$coefficients[i])/(1+exp(mod4$coefficients[i]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i])/(1+exp(mod4$coefficients[i]))
probacum2=exp(mod4$coefficients[i-1])/(1+exp(mod4$coefficients
[i-1]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1])/(1+exp(mod4$
coefficients[i-1]))
marginal=probacum-probacum2}
}}}

#(marca==1 & gosto==2 & categoria==i)
else{if(marca==1 & gosto==2 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[9])/
(1+exp (mod4$coefficients[i]-mod4$coefficients[9]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[9])/
(1+exp (mod4$coefficients[i]-mod4$coefficients[9]))

```

```

probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[9])/
(1+exp (mod4$coefficients[i-1]-mod4$coefficients[9]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[9])/
(1+exp (mod4$coefficients[i-1]-mod4$coefficients[9]))
marginal=probacum-probacum2}
}}

```

```

#(marca==1 & gosto==3 & categoria==i)
else{if(marca==1 & gosto==3 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[10])/
(1+exp (mod4$coefficients[i]-mod4$coefficients[10]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[10])/
(1+exp (mod4$coefficients[i]-mod4$coefficients[10]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[10])/
(1+exp (mod4$coefficients[i-1]-mod4$coefficients[10]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[10])/
(1+exp (mod4$coefficients[i-1]-mod4$coefficients[10]))
marginal=probacum-probacum2}
}}

```

```

#(marca==2 & gosto==1 & categoria==i)
else{if(marca==2 & gosto==1 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[5])/
(1+exp (mod4$coefficients[i]-mod4$coefficients[5]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[5])/

```



```

(1+exp (mod4$coefficients[i]-mod4$coefficients[5]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[5])/
(1+exp (mod4$coefficients[i-1]-mod4$coefficients[5]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[5])/
(1+exp (mod4$coefficients[i-1]-mod4$coefficients[5]))
marginal=probacum-probacum2}
}}}

#(marca==2 & gosto==2 & categoria==i)
else{if(marca==2 & gosto==2){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[5]-
mod4$coefficients[9])/ (1+exp(mod4$coefficients[i]-mod4$coefficients[5]
-mod4$coefficients[9]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[5]
-mod4$coefficients[9])/ (1+exp(mod4$coefficients[i]-mod4$coefficients[5]
-mod4$coefficients[9]))probacum2=exp(mod4$coefficients[i-1]-mod4$
coefficients[5]-mod4$coefficients [9])/(1+exp(mod4$coefficients[i-1]-
mod4$
coefficients[5]-mod4$coefficients[9]))marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[5]-mod4$
coefficients[9])/(1+exp(mod4$coefficients[i-1]-mod4$coefficients[5]
-mod4$coefficients[9]))
marginal=probacum-probacum2}
}}}}

#(marca==2 & gosto==3 & categoria==i)
else{if(marca==2 & gosto==3 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[5]-mod4$

```

```

coefficients[10])/ (1+exp(mod4$coefficients[i]-mod4$
coefficients[5]-mod4$coefficients[10]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[5]-mod4$
coefficients[10])/ (1+exp(mod4$coefficients[i]-mod4$
coefficients[5]-mod4$
coefficients[10]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[5]-
mod4$coefficients[10])/(1+exp(mod4$coefficients[i-1]-mod4$
coefficients[5]-mod4$coefficients [10]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[5]-mod4$
coefficients[10])/(1+exp(mod4$coefficients[i-1]-mod4$
coefficients[5]-mod4$coefficients [10]))
marginal=probacum-probacum2}
}}

#(marca==3 & gosto==1 & categoria==i)
else{if(marca==3 & gosto==1 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[6])/
(1+exp (mod4$coefficients[i]-mod4$coefficients[6]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[6])/
(1+exp (mod4$coefficients[i]-mod4$coefficients[6]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[6])
/(1+exp (mod4$coefficients[i-1]-mod4$coefficients[6]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[6])
/(1+exp (mod4$coefficients[i-1]-mod4$coefficients[6]))
marginal=probacum-probacum2}

```

```
}}}
```

```

#(marca==3 & gosto==2 & categoria==i)
else{if(marca==3 & gosto==2 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[6]-mod4$
coefficients[9])/ (1+exp(mod4$coefficients[i]-mod4$
coefficients[6]-mod4$coefficients[9]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[6]-mod4$
coefficients[9])/ (1+exp(mod4$coefficients[i]-mod4$
coefficients[6]-mod4$coefficients[9]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[6]
-mod4$coefficients [9])/(1+exp(mod4$coefficients[i-1]-
mod4$coefficients[6]-mod4$coefficients[9]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[6]-mod4$
coefficients [9])/(1+exp(mod4$coefficients[i-1]-mod4$
coefficients[6]-mod4$coefficients[9]))
marginal=probacum-probacum2}
}}}
```

```

#(marca==3 & gosto==3 & categoria==i)
else{if(marca==3 & gosto==3 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[6]-mod4$
coefficients[10])/(1+exp(mod4$coefficients[i]-mod4$
coefficients[6]-mod4$coefficients[10]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[6]-mod4$
coefficients[10])/(1+exp(mod4$coefficients[i]-mod4$
coefficients[6]-mod4$coefficients[10]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients
[6]-mod4$coefficients [10])/(1+exp(mod4$coefficients
```

```

[i-1]-mod4$coefficients[6]-mod4$coefficients [10]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[6]-mod4$
coefficients [10])/(1+exp(mod4$coefficients[i-1]-mod4$
coefficients[6]-mod4$coefficients [10]))
marginal=probacum-probacum2}
}}}

#(marca==4 & gosto==1 & categoria==i)
else{if(marca==4 & gosto==1 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[7])/(1+exp
(mod4$coefficients[i]-mod4$coefficients[7]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[7])/(1+exp
(mod4$coefficients[i]-mod4$coefficients[7]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[7])/(1+exp
(mod4$coefficients[i-1]-mod4$coefficients[7]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[7])/(1+exp
(mod4$coefficients[i-1]-mod4$coefficients[7]))
marginal=probacum-probacum2}
}}}

#(marca==4 & gosto==2 & categoria==i)
else{if(marca==4 & gosto==2 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[7]-mod4$
coefficients[9])/(1+exp(mod4$coefficients[i]-mod4$
coefficients[7]-mod4$coefficients[9]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[7]-mod4$

```

```

coefficients[9))/(1+exp(mod4$coefficients[i]-mod4$
coefficients[7]-mod4$coefficients[9]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients
[7]-mod4$coefficients [9))/(1+exp(mod4$coefficients[i-1]
-mod4$coefficients[7]-mod4$coefficients[9]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[7]-mod4$
coefficients [9))/(1+exp(mod4$coefficients[i-1]-mod4$
coefficients[7]-mod4$coefficients[9]))
marginal=probacum-probacum2}
}}}

#(marca==4 & gosto==3 & categoria==i)
else{if(marca==4 & gosto==3){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[7]-mod4$
coefficients[10))/(1+exp(mod4$coefficients[i]-mod4$
coefficients[5]-mod4$coefficients[8]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[7]-mod4$
coefficients[10))/(1+exp(mod4$coefficients[i]-mod4$
coefficients[7]-mod4$coefficients[10]))probacum2=exp(mod4$
coefficients[i-1]-mod4$coefficients[7]-mod4$coefficients [10])
/(1+exp(mod4$coefficients[i-1]-mod4$coefficients[7]-mod4$
coefficients [10]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[7]-
mod4$coefficients [10))/(1+exp(mod4$coefficients[i-1]-
mod4$coefficients[7]-mod4$coefficients [10]))
marginal=probacum-probacum2}
}}}

```

```

#(marca==5 & gosto==1 & categoria==i)
else{if(marca==5 & gosto==1 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[8])/(1+
exp(mod4$coefficients[i]-mod4$coefficients[8]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[8])/(1+exp
(mod4$coefficients[i]-mod4$coefficients[8]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[8])/(1+exp
(mod4$coefficients[i-1]-mod4$coefficients[8]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[8])/(1+exp
(mod4$coefficients[i-1]-mod4$coefficients[8]))
marginal=probacum-probacum2}
}}}}

```

```

#(marca==5 & gosto==2 & categoria==i)
else{if(marca==5 & gosto==2 ){if(i==1){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[8]-
mod4$coefficients[9])/(1+exp(mod4$coefficients[i]-mod4$
coefficients[8]-mod4$coefficients[9]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients[i]-mod4$coefficients[8]
-mod4$coefficients[9])/
(1+exp(mod4$coefficients[i]-mod4$coefficients[8]-mod4$
coefficients[9]))
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients
[8]-mod4$coefficients [9])/(1+exp(mod4$coefficients[i-1]
-mod4$coefficients[8]-mod4$coefficients[9]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients[i-1]-mod4$coefficients[8]-

```

```

mod4$coefficients [9])/
(1+exp(mod4$coefficients [i-1]-mod4$coefficients [8]-mod4$
coefficients [9]))
marginal=probacum-probacum2}
}}

#(marca==5 & gosto==3 & categoria==i)
else{if(marca==5 & gosto==3){if(i==1){
probacum=exp(mod4$coefficients [i]-mod4$coefficients [8]-mod4$
coefficients [10])/(1+exp(mod4$coefficients [i]-mod4$coefficients
[8]-mod4$coefficients [10]))
marginal=probacum}
else{if(i>1 & i<5){
probacum=exp(mod4$coefficients [i]-mod4$coefficients [8]-mod4$
coefficients [10])/(1+exp(mod4$coefficients [i]-mod4$
coefficients [8]-mod4$coefficients [10]))probacum2=
exp(mod4$coefficients [i-1]-mod4$coefficients [8]-
mod4$coefficients [10])/(1+exp(mod4$coefficients [i-1]
-mod4$coefficients [8]-mod4$coefficients [10]))
marginal=probacum-probacum2}
else{if(i==5){
probacum=1
probacum2=exp(mod4$coefficients [i-1]-mod4$coefficients [8]-
mod4$coefficients [10])/
(1+exp(mod4$coefficients [i-1]-mod4$coefficients [8]-mod4$
coefficients [10]))
marginal=probacum-probacum2}
}}}}
}}}}}}
}}}}}}
}}}}}}

marg=c("Probabilidade marginal"=marginal)
print(marg)

```

}

X11=c(marg(1,1,1),marg(2,1,1),marg(3,1,1),marg(4,1,1),
 ,marg(5,1,1),marg(1,2,1),marg(2,2,1),marg(3,2,1),
 marg(4,2,1),marg(5,2,1),marg(1,3,1),marg(2,3,1),
 marg(3,3,1),marg(4,3,1),marg(5,3,1))X22=c(marg(1,1,2),
 marg(2,1,2),marg(3,1,2),marg(4,1,2),marg(5,1,2),
 marg(1,2,2),marg(2,2,2),marg(3,2,2),marg(4,2,2),
 marg(5,2,2),marg(1,3,2),marg(2,3,2),marg(3,3,2),
 marg(4,3,2),marg(5,3,2))

X33=c(marg(1,1,3),marg(2,1,3),marg(3,1,3),marg(4,1,3),
 marg(5,1,3),marg(1,2,3),marg(2,2,3),marg(3,2,3),
 marg(4,2,3),marg(5,2,3),marg(1,3,3),marg(2,3,3),
 marg(3,3,3),marg(4,3,3),marg(5,3,3))

X44=c(marg(1,1,4),marg(2,1,4),marg(3,1,4),marg(4,1,4),
 ,marg(5,1,4),marg(1,2,4),marg(2,2,4),marg(3,2,4),
 marg(4,2,4),marg(5,2,4),marg(1,3,4),marg(2,3,4),
 marg(3,3,4),marg(4,3,4),marg(5,3,4))

X55=c(marg(1,1,5),marg(2,1,5),marg(3,1,5),marg(4,1,5),
 marg(5,1,5),marg(1,2,5),marg(2,2,5),marg(3,2,5),
 marg(4,2,5),marg(5,2,5),marg(1,3,5),marg(2,3,5),
 marg(3,3,5),marg(4,3,5),marg(5,3,5))

x11=c(0.3608196,0.277565,0.02732119,0.1721031,0.1780237,
 0.2318162,0.1703925,0.01479345,0.1000139,0.1037653,
 0.2656153,0.1975389,0.01767855,0.1160801,0.121856)

x22=c(0.4130473,0.4220599,0.118183,0.3854629,0.3896312,
 0.4147548,0.3841981,0.06864023,0.3025037,0.3086508,
 0.4211631,0.4012297,0.08069023,0.3291952,0.3350299)

x33=c(0.09656362,0.1209146,0.1050212,0.1545707,
 0.152813,0.1356225,0.1550689,0.06816873,0.1669081,
 0.167032,0.1246924,0.146746,0.07802295,0.1664271,
 0.1659514)

x44=c(0.1033781,0.1414456,0.3986011,0.2197982,
 0.214022,0.1699045,0.2215068,0.3456291,0.3103727,
 0.304619,0.1482424,0.1963915,0.3660127,0.2845135,


```
0.27854)
x55=c(0.02619139,0.03801491,0.3508736,0.06806512,
0.06551001,0.0479021,0.06883368,0.5027685,0.1202016,
0.1159328,0.04028688,0.05809389,0.4575956,0.1023279
,0.0986227)

pred.mat11=cbind(x11,x22,x33,x44,x55)

lab <- paste("Gosto=", rep(levels(gostodeoutrossuodefutas), 5), ", ",
"marca=", rep(levels(marca), each = 3), sep = "")
lab <- paste("Marca=", rep(levels(marca), 3), ", ",
"Outros=", rep(levels(gostodeoutrossuodefutas), each = 5), sep = "")

par(mfrow = c(3, 5))
for (k in 1:15) {
plot(1:5, pred.mat11[k,], lty = 1, type = "l", ylim = c(0,
1), xlab = "Impressão Global", axis = FALSE, ylab = "Probabilidades
estimadas",foont.ylab=7,
main = lab[ceiling(k)], las = 1)
#axis(1)
#axis(2)
}
```

ANEXO

Análise sensorial de suco de caju

Nome:

Sexo: () Feminino () Masculino

Faixa etária: () menos de 18 () 19 - 25 () 26 - 35

() 36 - 45 () 46 - 55 () mais de 56

1 .

Categoria	Suco de caju	Outros sucos de frutas
Consumo muito frequentemente	()	()
Consumo muito	()	()
Consumo ocasionalmente	()	()
Consumo pouco	()	()
Não consumo	()	()

2 .

Categoria	Suco de caju	Outros sucos de frutas
Gosto muitíssimo	()	()
Gosto muito	()	()
Gosto moderadamente	()	()

3. Por favor, utilizando a escala abaixo, descreva o quanto você gostou ou desgostou de cada amostra com relação à: COR, HOMOGENEIDADE, AROMA, SABOR, CORPO (VISCOSIDADE) E IMPRESSÃO GLOBAL.

ESCALA:

1. Gostei muitíssimo
2. Gostei muito
3. Gostei moderadamente
4. Gostei ligeiramente
5. Nem gostei, nem desgostei
6. Desgostei ligeiramente
7. Desgostei moderadamente
8. Desgostei muito
9. Desgostei muitíssimo

Amostra	Cor	Homogeneidade	Aroma	Sabor	Corpo	Impressão Global

4. Descreva o que você achou POSITIVO e/ ou NEGATIVO em cada amostra:

Amostra	Positivo	Negativo

5. Avalie o QUÃO IDEAL você acha a INTENSIDADE de DOÇURA, CORPO (VIS-COSIDADE), ACIDEZ e ADSTRINGÊNCIA de cada amostra codificada, utilizando a escala abaixo:

Categoria	Amostra	Doçura	Corpo	Acidez	Adstringência
+4 Extremamente mais forte que o ideal					
+3 Muito mais forte que o ideal					
+2 Moderadamente mais forte que o ideal					
+1 Ligeiramente mais forte que o ideal					
0 Ideal					
-1 Ligeiramente menos forte que o ideal					
-2 Moderadamente menos forte que o ideal					
-3 Muito menos forte que o ideal					
-4 Extremamente menos forte que o ideal					

6. Assinale para cada uma das amostras, qual seria a sua atitude de consumo do produto usando a escala abaixo:

Escala	A	B	C	D	E
Certamente consumiria					
Provavelmente consumiria					
Tenho dúvidas se consumiria					
Provavelmente não consumiria					
Certamente na consumiria					