

CHOICE-BASED CONJOINT ANALYSIS: UM ENFOQUE BAYESIANO

Eduardo Campana BARBOSA¹
Carlos Henrique Osório SILVA¹
Moysés NASCIMENTO¹
Fabyano Fonseca e SILVA²
Valéria Paula Rodrigues MINIM³
Rosires DELIZA⁴
Suzana Maria DELLA LUCIA⁵

- RESUMO: Neste trabalho apresenta-se um enfoque Bayesiano para o modelo Frequentista Logit Multinomial empregado na *Choice-Based Conjoint Analysis*. A análise foi conduzida com dados de escolha referentes à avaliação, por 144 consumidores, de oito amostras de iogurtes *light* sabor morango, variando a informação de três ingredientes (açúcar, gordura e proteína) em um delineamento fatorial completo. São apresentados os resultados e inferências obtidos pelo enfoque Bayesiano, no que se refere à estimação do efeito principal dos atributos, das probabilidades e da razão de escolhas. Os resultados Frequentistas também são informados e discutidos. A análise Bayesiana apresentou resultados similares aos Frequentistas, no entanto, permitiu a construção de intervalos de credibilidade para probabilidades e razões de escolha, possibilitando comparar estatisticamente tais quantidades. Quanto aos resultados práticos, a maior probabilidade de escolha foi associada ao iogurte *light* sabor morango contendo as informações “0% de açúcar”, “0% de gordura” e “enriquecido com proteínas bioativas”.
- PALAVRAS-CHAVE: Modelo Logit Multinomial; fatorial; escolha por iogurtes.

1 Introdução

Muitos fatores contribuem para que o consumidor faça uma escolha no momento da compra. Dentre estes, merece especial destaque a qualidade e o preço do produto ofertado. Adicionalmente, características intrínsecas como a marca, o design, a embalagem e as informações presentes no rótulo podem interferir nesta decisão, fato explicado devido às expectativas, crenças ou o conhecimento prévio dos consumidores em relação a cada produto (DELIZA *et al.*, 2003).

¹ Universidade Federal de Viçosa - UFV, Departamento de Estatística, Viçosa, MG, Brasil. E-mail: duducampana@hotmail.com; moysesnascim@ufv.br; chos@ufv.br

² Universidade Federal de Viçosa - UFV, Departamento de Zootecnia, Viçosa, MG, Brasil. E-mail: fabyanofonseca@ufv.br

³ Universidade Federal de Viçosa - UFV, Departamento de Tecnologia de Alimentos, Viçosa, MG, Brasil. E-mail: vprm@ufv.br

⁴ Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA Agroindústria de Alimentos, Rio de Janeiro, RJ, Brasil. E-mail: rosires.deliza@embrapa.br

⁵ Universidade Federal do Espírito Santo - UFV, Departamento de Engenharia de Alimentos, Centro de Ciências Agrárias e Engenharias, Alegre, ES, Brasil. E-mail: smllucia@yahoo.com.br

Naturalmente, o sucesso e a sobrevivência de uma empresa estão atrelados à sua capacidade de fidelizar clientes. O bom relacionamento e a comunicação efetiva são princípios importantes para se atingir este objetivo, no entanto, não são suficientes. Logo, estratégias de *marketing* são desenvolvidas para se tentar compreender o comportamento e a preferência dos consumidores no momento da compra, visando maximizar os lucros e minimizar os custos e riscos (NATTER e FEURSTEIN, 2002).

É neste contexto que a *Choice-Based Conjoint Analysis* (CBCA) é aplicada. Tal metodologia permite modelar a escolha do consumidor e estimar probabilidades associadas às diferentes alternativas de produtos. Moore (2004) observou que a modelagem de escolhas é uma forma realista, simples e prática para avaliar a preferência do consumidor. Na CBCA, as variáveis em estudo são atributos específicos dos produtos, definidos por seus respectivos níveis, isto é, valores ou características que os quantificam ou qualificam (AAKER, KUMAR e DAY, 2008).

As aplicações clássicas da CBCA englobam o estabelecimento de novas estratégias competitivas (MAHAJAN, GREEN e GOLDBERG, 1982), a avaliação de novos produtos (LOUVIERE e WOODWORTH, 1983) e a otimização de produtos já existentes (GREEN e KRIEGER, 1993). Trabalhos mais recentes, em diferentes áreas do conhecimento, podem também ser verificados em Tempesta *et al.* (2010), Zimmermann *et al.* (2013), Utz *et al.* (2014) e Arenoe *et al.* (2015). No Brasil, a referida técnica ainda é pouco explorada. No entanto, na área de alimentos, Deliza *et al.* (2010) utilizaram a CBCA para avaliar o efeito do preço, da informação sobre irradiação e da qualidade sensorial na escolha de mamão e Della Lucia *et al.* (2010) utilizaram a CBCA para avaliar a influência de alguns atributos presentes na embalagem de iogurte *light* sabor morango na preferência do consumidor.

Com o avanço dos métodos computacionais, uma alternativa de análise estatística que tem recebido destaque é a inferência Bayesiana. Nesse enfoque, o parâmetro não é tratado como fixo e desconhecido, e sim como uma variável aleatória associada a um modelo de probabilidade. Portanto, infere-se sobre este através de sua distribuição, denominada como distribuição *a posteriori*, que é obtida pela combinação da informação presente nos dados amostrais, via função de verossimilhança, com a informação denominada de *a priori*, que é a distribuição de probabilidade do parâmetro definida pelo pesquisador, considerada como a informação que ele possui sobre o mesmo (BAYES e PRICE, 1763).

Rossi *et al.* (1996), Allenby e Rossi (1999) e Rossi *et al.* (2005) foram os principais responsáveis por introduzir o enfoque Bayesiano para a classe dos modelos de escolha discreta (GREENE, 2003; TRAIN, 2009). Mais especificamente, essa abordagem permite que o procedimento de estimação dos parâmetros do modelo Logit Multinomial, empregado na CBCA, não envolva restrições quanto a graus de liberdade (BARRETO, 2004) e não exija a maximização de uma função específica (TRAIN, 2009), o que é uma vantagem, uma vez que o algoritmo iterativo falha em convergir por problemas de valores iniciais errôneos ou inexistência de um máximo global da função objetivo, principalmente em pequenas amostras, isto é, quando o número de consumidores avaliados não é suficientemente grande.

Um aspecto interessante sobre o enfoque proposto está associado aos intervalos de confiança Bayesianos ou intervalos de credibilidade para os parâmetros de interesse, obtidos pelos *quantis* das respectivas distribuições *a posteriori*. Logo, torna-se possível testar hipóteses ou realizar comparações estatísticas sobre os parâmetros ou qualquer função real destes, o que na inferência clássica nem sempre é trivial, já que a distribuição amostral de uma determinada estatística pode não ser conhecida.

Neste sentido, expõe-se aqui uma aplicação do enfoque Bayesiano na CBCA. Além dos resultados tradicionais, será possível explorar a distribuição *a posteriori* das probabilidades e das razões de escolha, bem como inferir sobre tais quantidades, sendo esse o principal objetivo do presente trabalho, uma vez que no enfoque Frequentista isso não é realizado. Os dados utilizados são escolhas de consumidores que avaliaram a influência de três atributos na embalagem de iogurte *light* sabor morango. Adicionalmente, conforme mencionado por Daziano (2013), poucos trabalhos na literatura ilustram o enfoque Bayesiano aplicado ao modelo da CBCA. Portanto, tem-se também como objetivo descrever as principais vantagens e apresentar uma nova alternativa metodológica para a análise de dados dessa natureza.

2 Material e métodos

2.1 Produtos avaliados. Descrição do conjunto de dados

Para avaliar a escolha dos consumidores definiu-se $R = 3$ atributos: informação sobre o conteúdo de açúcar (“Açúcar”), informação sobre o conteúdo de gordura (“Gordura”) e informação sobre o conteúdo de proteína (“Proteína”), com os respectivos níveis: açúcar (“0% de açúcar” e “Com adoçante”), gordura (“0% de gordura” e “Baixo teor de gordura”) e proteína (“Com proteínas do soro do leite” e “Com proteínas bioativas”). O delineamento fatorial empregado foi o do perfil completo, com $J = 2^3 = 8$ amostras de iogurte, apresentados na Tabela 1. Os níveis dos atributos foram codificados por variáveis *dummy*.

Tabela 1 – Delineamento Experimental do estudo

Amostra	Açúcar	Gordura	Proteína
1	0% de açúcar	0% de gordura	Com proteínas do soro do leite
2	Com adoçante	0% de gordura	Com proteínas do soro do leite
3	0% de açúcar	Baixo teor de gordura	Com proteínas do soro do leite
4	Com adoçante	Baixo teor de gordura	Com proteínas do soro do leite
5	0% de açúcar	0% de gordura	Com proteínas bioativas
6	Com adoçante	0% de gordura	Com proteínas bioativas
7	0% de açúcar	Baixo teor de gordura	Com proteínas bioativas
8	Com adoçante	Baixo teor de gordura	Com proteínas bioativas

O estudo envolveu $N = 144$ consumidores que residiam na cidade de Viçosa (Minas Gerais, Brasil) e o pré-requisito para que o voluntário participasse da pesquisa era que ele tivesse o hábito de consumir iogurte e produtos *light*. Destes, 59,7% foram do sexo feminino e 88,2% encontravam-se com idade entre 15 e 25 anos; 81,2% possuíam curso superior incompleto ou em andamento. Adicionalmente, 40,3% dos participantes afirmaram ter renda entre 1 e 5 salários mínimos; 38,2% entre 6 e 10 salários; 19,4% possuíam renda entre 11 e 20 salários e 2,1% mencionaram uma renda acima de 20 salários mínimos.

A avaliação foi conduzida em um local confortável e sob a orientação dos aplicadores do estudo. As oito amostras de iogurte foram apresentadas aos consumidores por meio de fotografias das respectivas embalagens. A ordem de apresentação foi balanceada segundo MacFie *et al.* (1989), no intuito de permitir a estimação não tendenciosa dos efeitos principais dos atributos e eliminar os efeitos da ordem de apresentação e residual, caracterizados pela influência de uma embalagem na avaliação da subsequente. A variável resposta foi codificada da seguinte forma: para a j -ésima amostra de iogurte escolhida pelo n -ésimo consumidor foi atribuído o valor 1 (um) e para as demais, não escolhidas, o valor 0 (zero). Maiores detalhes sobre os procedimentos descritos na seção 2.1 podem ser consultados em Della Lúcia *et al.* (2010).

2.2 Análise frequentista dos dados

A utilidade aleatória representa o benefício ou a satisfação do consumidor ao escolher uma determinada amostra de iogurte. Este termo é composto por um componente fixo e observável e outro aleatório, portanto, não observável. O modelo para a utilidade aleatória é apresentado como

$$U = X\beta + \varepsilon, \quad (1)$$

em que:

$U = (U_{11} \dots U_{1J} U_{21} \dots U_{2J} \dots U_{N1} \dots U_{NJ})'$ é o vetor ($NJ \times 1$) de utilidades aleatórias associadas aos J tratamentos, atribuídas pelos N indivíduos;

$X = [X_1 \ X_2 \ \dots \ X_N]'$, em que $X_1 = X_2 = \dots = X_N$ são matrizes ($J \times R$) que especificam a codificação do s -ésimo nível do r -ésimo atributo, presente no j -ésimo tratamento. Para $n = 1, 2, \dots, N$ tem-se que:

$$X_n = \begin{bmatrix} X_1^1 & X_2^1 & \dots & X_R^1 \\ X_1^2 & X_2^2 & \dots & X_R^2 \\ \vdots & \vdots & \dots & \vdots \\ X_1^J & X_2^J & \dots & X_R^J \end{bmatrix}$$

$\beta = (\beta_1 \ \beta_2 \ \dots \ \beta_R)'$ é o vetor de parâmetros desconhecidos (ou coeficientes de preferência) de dimensão ($R \times 1$), em que β_r representa o efeito do r -ésimo atributo sobre a utilidade aleatória média.

$\varepsilon = (\varepsilon_{11} \dots \varepsilon_{1J} \ \varepsilon_{21} \dots \varepsilon_{2J} \ \dots \ \varepsilon_{N1} \dots \varepsilon_{NJ})'$ é o vetor ($NJ \times 1$) de erros aleatórios e não observáveis do modelo, cuja suposição é de que cada ε_{nj} seja independente e possua distribuição de valores extremos do tipo I ou Gumbel, com média nula e variância constante e igual a $\frac{\pi^2}{6}$, sendo este valor definido para assegurar que o modelo (2) seja identificável (TRAIN, 2009).

O modelo (2) ou Logit Multinomial (McFADDEN, 1974) é derivado supondo que o n -ésimo consumidor escolherá a j -ésima amostra de iogurte, se e somente se, $U_{nj} > U_{nk}, \forall j \neq k; j, k \in \{1, 2, \dots, J\}$. Portanto,

$$P(Y_n = j|\mathbf{X}) = \frac{e^{x'_j\boldsymbol{\beta}}}{\sum_{k=1}^J e^{x'_k\boldsymbol{\beta}}} \quad \forall j = 1, 2, \dots, J, \quad (2)$$

em que $P(Y_n = j|\mathbf{X})$ é a probabilidade do n -ésimo consumidor escolher a j -ésima amostra de iogurte, condicional ao conjunto de atributos \mathbf{X} . Como a matriz \mathbf{X} é a mesma para $n = 1, 2, \dots, N$, tem-se que $P(Y_n = j|\mathbf{X}) = P(Y = j|\mathbf{X})$. Os estimadores do vetor $\boldsymbol{\beta} = (\beta_1 \beta_2 \dots \beta_R)'$ foram obtidos por Máxima Verossimilhança Parcial (SO e KUHFIELD, 1995; AIZAKI e NISHIMURA, 2008) e maximizam $L(\boldsymbol{\beta}, \mathbf{y}|\mathbf{X})$, definida em (3) como:

$$L(\boldsymbol{\beta}, \mathbf{y}|\mathbf{X}) = \prod_{n=1}^N \prod_{j=1}^J \left(\frac{e^{x'_j\boldsymbol{\beta}}}{\sum_{k=1}^J e^{x'_k\boldsymbol{\beta}}} \right)^{I_{nj}}. \quad (3)$$

Como a solução do sistema de equações diferenciais não é analítica emprega-se o método numérico de Newton-Raphson (GALLANT, 2009).

Na equação (3), $I_{nj} = 1$ se o n -ésimo consumidor escolheu a j -ésima amostra de iogurte e $I_{nj} = 0$ caso contrário. Para detectar quais atributos influenciaram a escolha do consumidor o teste de Wald (1943) para os coeficientes estimados utilizado. A importância dos níveis de cada atributo foi avaliada pela Razão de Escolhas (RE) (GREENE, 2003), definida pelo quociente entre duas probabilidades de escolha, conforme em (4):

$$RE(X_p, X_q) = \frac{P(Y = p|\mathbf{X})}{P(Y = q|\mathbf{X})} = e^{(x'_p - x'_q)\boldsymbol{\beta}}. \quad (4)$$

Por esta relação, é possível concluir o quanto o nível p de um atributo é preferível ($RE > 1$) ou não ($RE < 1$) em relação a outro nível q , dado que os níveis dos demais atributos não variam. Adicionalmente, se $RE = 1$ os níveis p e q de um determinado atributo são igualmente preferíveis. Destaca-se que a RE é constante e depende exclusivamente de características das alternativas p e q . Por isso, a remoção ou a inclusão de alternativas de escolhas não a altera. Essa característica é particular do modelo Logit Multinomial, devido à sua propriedade I.I.A. (*Independence of Irrelevant Alternatives*), assegurada principalmente pela independência dos erros aleatórios (TRAIN, 2009).

2.3 Análise bayesiana dos dados

A distribuição conjunta *a posteriori* dos parâmetros do modelo Logit Multinomial, $\pi(\boldsymbol{\beta}|\mathbf{y}, \mathbf{X})$, é dada por:

$$\pi(\boldsymbol{\beta}|\mathbf{y}, \mathbf{X}) = \frac{L(\mathbf{y}, \mathbf{X}|\boldsymbol{\beta}) p(\boldsymbol{\beta})}{\pi(\mathbf{y}, \mathbf{X})}$$

$$\pi(\boldsymbol{\beta}|\mathbf{y}, \mathbf{X}) \propto \left\{ \prod_{n=1}^N \prod_{j=1}^J \left(\frac{e^{x'_{nj}\boldsymbol{\beta}}}{\sum_{k=1}^J e^{x'_{nk}\boldsymbol{\beta}}} \right)^{I_{nj}} \right\} p(\boldsymbol{\beta}),$$

em que a função de verossimilhança, $L(\mathbf{y}, \mathbf{X}|\boldsymbol{\beta})$, possui as informações contidas na amostra aleatória e $p(\boldsymbol{\beta})$ é a distribuição conjunta *a priori*, que modela a incerteza do pesquisador associada ao vetor de parâmetros de interesse. A distribuição preditiva, $\pi(\mathbf{y}, \mathbf{X})$, não é indexada por $\boldsymbol{\beta}$ e por isso é incorporada à proporção.

Koop e Poirier (1993) verificaram que o modelo Logit Multinomial, empregado na CBCA, não apresenta conjugação natural. Neste sentido, adotou-se a sugestão proposta por Rossi *et al.* (2005), de uma distribuição *a priori* normal R -variada, com função densidade de probabilidade apresentada em (5) e a aproximação de Laplace (1986) para a função de verossimilhança conforme em (6):

$$p(\boldsymbol{\beta}) \propto |A|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} [(\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})' A (\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})] \right\}, \quad (5)$$

$$L(\mathbf{y}, \mathbf{X}|\boldsymbol{\beta}) \propto \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' H (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \right\}, \quad (6)$$

em que $\bar{\boldsymbol{\beta}}$ e A são, respectivamente, o vetor ($R \times 1$) de valores médios e a matriz ($R \times R$) de precisão ($A = \Sigma^{-1}$) da distribuição conjunta *a priori*. H é a matriz ($R \times R$) de precisão ou informação de Fisher observada, representada pelo o negativo da matriz Hessiana avaliada na moda *a posteriori* ou $\hat{\boldsymbol{\beta}}$, ou seja, $H = - \left\{ \frac{\partial^2 \ln[L(\boldsymbol{\beta}, \mathbf{y}|\mathbf{X})]}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \right\} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}}$. Para verificar os resultados do enfoque Bayesiano, na ausência de conhecimento prévio do pesquisador, por essa ser uma situação prática comum, optou-se por uma distribuição *a priori* pouco informativa (*flat*), com vetor de médias nulo e baixa precisão, isto é $\boldsymbol{\beta} \sim N_R(\mathbf{0}; 0.01\mathbf{I})$, em que \mathbf{I} é uma matriz ($R \times R$) identidade.

Com essa especificação a combinação de (5) e (6) implica em uma conjugação natural de distribuições normais R -variadas. Portanto, uma distribuição conjunta *a posteriori* com o núcleo de uma densidade normal R -variada, que com a seleção dos valores paramétricos para a distribuição *a priori* de $\boldsymbol{\beta}$, pode ser verificada em (7):

$$\pi(\boldsymbol{\beta}|\mathbf{y}, \mathbf{X}) \propto |H|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' H (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \right\} \quad (7)$$

As amostras das distribuições *a posteriori* dos parâmetros foram geradas pelo método de simulação estocástica conhecido como *Markov Chain Monte Carlo* (MCMC), via algoritmo *Independence Metropolis Hasting*. A distribuição *t-Student* R -

variada com $v = 6$ graus de liberdade, vetor de médias $\boldsymbol{\mu} = \widehat{\boldsymbol{\beta}}$ e matriz de precisão $A = H$ foi adotada como distribuição de transição ou candidata (ZELLNER e ROSSI, 1984; ROSSI *et al.*, 2005). A função densidade de probabilidade dessa distribuição é apresentada em (8).

$$q(\mathbf{B}|v, \widehat{\boldsymbol{\beta}}, H) \propto |H|^{\frac{1}{2}} \left[v + (\mathbf{B} - \widehat{\boldsymbol{\beta}})' H (\mathbf{B} - \widehat{\boldsymbol{\beta}}) \right]^{-\frac{R+v}{2}} \quad (8)$$

O procedimento de simulação consistiu em gerar um vetor de valores \mathbf{B} a partir da distribuição de transição (8). Cada valor deste vetor foi então comparado com os respectivos valores do vetor $\boldsymbol{\beta}$. A aceitação de \mathbf{B} , em relação à $\boldsymbol{\beta}$, ocorreu com probabilidade $\alpha(\boldsymbol{\beta}, \mathbf{B})$, conforme apresentado em (9). Este procedimento foi incrementado e repetido um número de vezes suficientemente grande, no intuito de assegurar que a distribuição final de cada parâmetro convergisse para a distribuição de equilíbrio ou de interesse (GAMERMAN e LOPES, 2006). Naturalmente, foi preciso definir os valores iniciais para o vetor $\boldsymbol{\beta}$, que neste trabalho foram às estimativas Frequentistas de Máxima Verossimilhança Parcial. Para as demais iterações do algoritmo, os valores de \mathbf{B} foram comparados com valores gerados por uma distribuição uniforme contínua com suporte entre zero e um.

$$\alpha(\boldsymbol{\beta}, \mathbf{B}) = \min \left[1, \frac{\pi(\mathbf{B}) q(\boldsymbol{\beta})}{\pi(\boldsymbol{\beta}) q(\mathbf{B})} \right] \quad (9)$$

Adotou-se a média a *posteriori* como o estimador Bayesiano pontual, que minimiza a função risco ou de perda quadrática. Detalhes sobre esta e outras funções de perda simétricas, como a absoluta e a 0-1, podem ser consultados, respectivamente, em Doob (1949), LeCam (1958) e Yuan e Gooijer (2014). Uma segunda medida analisada foi o Erro de Monte Carlo (E.M.C.), que segundo Sorensen *et al.* (1994), representa a variância das estimativas em relação a dimensão da cadeia de Markov. O E.M.C. será pequeno se o seu valor, somado à estimativa Bayesiana do parâmetro, não resultar em alteração até a segunda casa decimal. Testes de hipóteses e comparações sobre os parâmetros foram conduzidos via intervalos H.P.D. (*Highest Posterior Density*), que possuem o menor comprimento e a maior densidade a *posteriori* (CHEN e SHAO, 1999).

2.4 Aspectos Computacionais

Utilizou-se o *software* R (R CORE TEAM, 2017) na condução das análises estatísticas. O procedimento de estimação Frequentista ocorreu pela função *clogit* do pacote *survival*. O algoritmo *Independence Metropolis Hasting* foi implementado pela função *rmnlIndepMetrop* do pacote *bayesm*. A convergência das cadeias foi avaliada graficamente e pelos critérios estatísticos de Geweke (1989) e Raftery e Lewis (1992), ambos disponíveis no pacote BOA (*Bayesian Output Analysis*) do *software* R. Para a inicialização do MCMC foram utilizadas 20.000 iterações, *burn-in* de 1.000 e *thin* de 2. O nível de significância adotado em análises e testes estatísticos foi de $\alpha = 5\%$.

3 Resultados e discussão

3.1 Estimação do vetor de parâmetros

A Tabela 2 apresenta as estimativas pontuais dos parâmetros com os seus respectivos erros padrão. O teste de Wald e a significância estatística de cada coeficiente também são informados.

Tabela 2 – Estimativa, erro padrão e teste de Wald para os coeficientes do modelo

Atributos	$\hat{\beta}$	Erro Padrão	Z	Valor-p
Açúcar	-1,6603	0,2275	-7,299	0,0000*
Proteína	1,5126	0,2166	6,982	0,0000*
Gordura	-0,7885	0,1798	-4,386	0,0000*

* Significativo a 5% de probabilidade

Pelo teste de Wald, a hipótese nula $H_0: \beta_i = 0$ foi rejeitada para os três atributos avaliados (valor-p < 0,05). Portanto, concluiu-se que informações referentes ao conteúdo de açúcar, proteína e gordura, presentes na embalagem do iogurte *light* sabor morango influenciaram significativamente a escolha dos consumidores.

Na Tabela 3 estão apresentados os resultados obtidos pelo método Bayesiano de estimação. Os valores exibidos foram sumarizados a partir das respectivas distribuições *a posteriori* dos parâmetros.

Tabela 3 – Média, Mediana (quantil de 50%), Moda, Desvio Padrão *a posteriori* (D.P.), Erro de Monte Carlo (E.M.C.) e intervalos H.P.D. com 95% de credibilidade

Atributos	Média	Mediana	Moda	D.P.	E.M.C.	H.P.D.	
						$\beta_{2,5\%}$	$\beta_{97,5\%}$
Açúcar	-1,6767	-1,6708	-1,6354	0,2296	0,0025	-2,1254	-1,2382
Proteína	1,5260	1,5221	1,4837	0,2202	0,0027	1,1090	1,9576
Gordura	-0,7960	-0,7904	-0,7815	0,1792	0,0019	-1,1461	-0,4539

As estimativas de média *a posteriori* foram semelhantes às obtidas pelo método de estimação Frequentista. Este é um resultado esperado, devido à combinação entre a aproximação de Laplace, desenvolvida em torno da moda *a posteriori* ($\hat{\beta}$), e a distribuição *a priori* pouco informativa ou *flat*. Logo, a média *a posteriori* converge para o estimador de máxima verossimilhança, neste caso, de máxima verossimilhança parcial (GAMERMAN e LOPES, 2006). Adicionalmente, como a média, a mediana e a moda *a posteriori* conduziram a uma mesma estimativa, devido à simetria e a uni-modalidade da distribuição *a posteriori* normal, o estimador Bayesiano minimizou não apenas a perda quadrática, mas também a perda absoluta e a perda 0-1. Portanto, a escolha desse estimador não penalizou o erro de estimação, mesmo se o seu emprego não fosse o mais adequado.

Ainda conforme a Tabela 3, o desvio padrão *a posteriori* se aproximou das estimativas Frequentistas de erro padrão e a análise dos intervalos H.P.D. também conduziu a conclusões similares sobre a significância dos coeficientes estimados. Os

baixos Erros de Monte Carlo (E.M.C.) informaram que a variância das estimativas Bayesianas não varia de forma significativa com a alteração da dimensão da cadeia, o que é um resultado desejável. Na Figura 1 estão apresentadas as cadeias de Markov (gráficos à esquerda) para os parâmetros de interesse e as respectivas funções de autocorrelação (gráficos à direita). Adicionalmente, as densidades estimadas podem ser consultadas no ANEXO I deste trabalho.

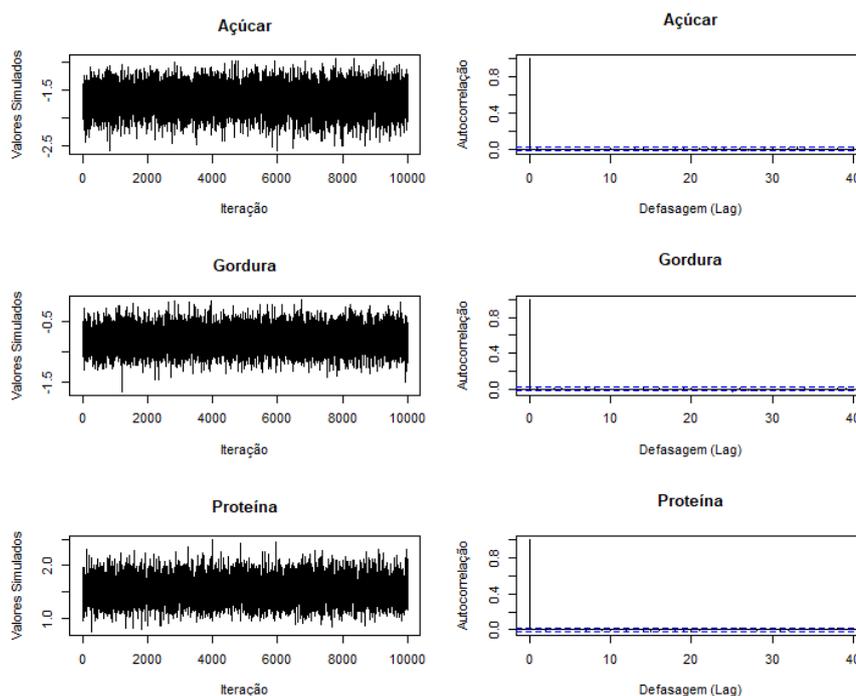


Figura 1 – Convergência das cadeias e função de autocorrelação.

De acordo com os gráficos da simulação a convergência foi alcançada, pois o comportamento das cadeias oscilou em torno de um valor médio constante, refletindo uma forma de equilíbrio ao longo das iterações. Como esperado, uma vez que o algoritmo *Independence Metropolis Hasting* gera os valores atuais da cadeia de forma independente dos valores passados, as funções de autocorrelação indicaram uma estrutura de dependência não significativa, pois os limites de confiança, estabelecidos a 5% de probabilidade, não foram ultrapassados. Por esse motivo, Rossi *et al.* (2005) e Daziano (2013), concluíram que o algoritmo *Independence Metropolis Hasting* foi mais eficiente, no sentido de induzir menor grau de autocorrelação e tempo computacional, do que os algoritmos *Random-Walk Metropolis Hasting* e *Slice-Sampler*, respectivamente. A Tabela 4 apresenta os resultados para os critérios de convergência de Raftery Lewis e de Geweke.

Tabela 4 – Resultados para os Critérios de Raftery-Lewis e Geweke

Atributos	Raftery-Lewis				Geweke	
	<i>thin</i>	<i>burn-in</i>	Total	Fator de Dependência	<i>z</i>	<i>valor-p</i>
Fat	1	2	3394	1,0662	0,9196	0,3577
Protein	1	2	3929	1,0485	0,6486	0,5165
Sugar	1	3	4095	1,0931	-0,4437	0,6572

O critério de Raftery-Lewis demonstrou que com menos de 5.000 iterações as distribuições simuladas já convergiram para as distribuições de equilíbrio ou de interesse, pois o Fator de Dependência foi menor do que cinco. Esse critério estimou ainda os valores necessários de *thin* e *burn-in*, que devido ao algoritmo utilizado foram muito baixos. O critério de Geweke corroborou a conclusão de convergência, visto que o valor- $p > 0,05$ para os três atributos. Tierney (1994) complementa que a adoção de uma distribuição *a priori* normal já oferece boas propriedades de convergência para muitos algoritmos MCMC. Adicionalmente, com a aproximação de Laplace para a função de verossimilhança, a densidade *a posteriori* possui caudas exponenciais, o que assegura algumas propriedades desejáveis, como a ergodicidade uniforme e geométrica das cadeias (ROSSI *et al.*, 2005).

3.2 Probabilidades de escolha

A Tabela 5 apresenta as probabilidades de escolha para as oito amostras ($j = 1, 2, \dots, 8$) de iogurte *light* sabor morango. Na análise Frequentista, tais quantidades foram estimadas via modelo (2), substituindo os valores de β pelos valores de $\hat{\beta}$. No enfoque Bayesiano, as cadeias simuladas para os parâmetros do modelo (2) foram utilizadas para construção das distribuições *a posteriori* das probabilidades de escolha. A média *a posteriori* foi também definida como estimador Bayesiano e é apresentada, bem como o desvio padrão *a posteriori* (D.P.) e os intervalos H.P.D. com 95% de credibilidade.

Tabela 5 – Probabilidades de Escolha Frequentistas e Bayesianas para as amostras de iogurte

Amostras de iogurte	$P(Y = j X)$	Enfoque Bayesiano			
		Média	D.P.	H.P.D.	
				$P_{j,2.5\%}$	$P_{j,97.5\%}$
1	0,1043	0,1045	0,0198	0,0661	0,1426
2	0,0198	0,0199	0,0054	0,0103	0,0308
3	0,0474	0,0475	0,0105	0,0277	0,0681
4	0,0090	0,0090	0,0027	0,0043	0,0144
5	0,4734	0,4733	0,0373	0,3997	0,5447
6	0,0900	0,0899	0,0182	0,0555	0,1270
7	0,2152	0,2150	0,0290	0,1596	0,2731
8	0,0409	0,0409	0,0095	0,0235	0,0602
Total	1,0000	1,0000	-	-	-

Pela análise frequentista concluiu-se que a amostra de iogurte com maior probabilidade de escolha, de acordo com os consumidores avaliados, foi a amostra 5, com probabilidade 0,4734 e características: 0% de açúcar, 0% de gordura e com proteínas bioativas. As amostras de iogurte 7 (0% de açúcar, baixo teor de gordura e enriquecido com proteínas bioativas) e 1 (0% de açúcar, 0% de gordura e enriquecido com proteínas de soro do leite) foram, respectivamente, a segunda (0,2152) e a terceira (0,1043) com as maiores probabilidades de escolha. A amostra de iogurte 4, com adoçante, baixo teor de gordura e proteínas do soro de leite, foi a que apresentou a menor probabilidade de escolha, igual a 0,0090.

Note que no enfoque Frequentista, apenas as estimativas pontuais de probabilidade de escolha são obtidas, enquanto no enfoque Bayesiano, devido ao conhecimento das distribuições a *posteriori* (densidades estimadas em ANEXO II), é possível obter uma medida de precisão sobre estas estimativas, tomando-se simplesmente o desvio padrão das amostras simuladas. Adicionalmente, com os intervalos H.P.D. tornou-se possível realizar comparações estatísticas entre as probabilidades de escolha. Logo, concluiu-se que a probabilidade estimada para a amostra 5 foi estatisticamente superior às probabilidades estimadas para as demais amostras de iogurte, pois não houve sobreposição dos limites de seu intervalo com os limites dos intervalos de nenhuma outra amostra. Para a amostra 7, que possuiu a segunda maior probabilidade de escolha, a conclusão foi similar. Já para as amostras de iogurte 1, 6 e 3, as probabilidades de escolha não diferiram estatisticamente entre si, uma vez que os respectivos intervalos não se sobrepõem. Com este resultado, um empresário/empreendedor pode optar por investir na amostra de iogurte que lhe oferecesse, por exemplo, menor custo de fabricação ou menor valor de compra, o que implicaria, principalmente, em redução de custos.

3.3 Razão de escolha

A Tabela 6 apresenta os resultados obtidos para a Razão de Escolha. Os valores informados foram calculados para superar a unidade, a fim de padronizar as interpretações. O estimador Bayesiano foi também a média a *posteriori* ou o valor esperado dessas distribuições, obtidas pela razão de duas distribuições das probabilidades de escolha, isto é, $RE = E[P_j/P_k]$, com $j \neq k$ e $j, k = 1, 2, \dots, 8$. Adicionalmente, o desvio padrão a *posteriori* (D.P.) e os intervalos H.P.D. com 95% de credibilidade também são apresentados.

Tabela 6 – Razão de Escolha para os atributos Açúcar, Gordura e Proteína

Razão de Escolha	$RE(X_p, X_q)$	Enfoque Bayesiano			
		Média	D.P.	H.P.D.	
				$RE_{2.5\%}$	$RE_{97.5\%}$
Açúcar	5,2609	5,5241	1,3260	3,2979	8,2554
Gordura	2,2000	2,2485	0,4099	1,5231	3,0763
Proteína	4,5385	4,7191	1,0621	2,8982	6,9042

Sob o aspecto sensorial, os resultados demonstraram que a amostra de iogurte com “0% de açúcar” foi 5,2609 vezes mais provável de ser escolhida do que a amostra de iogurte “Com adoçante”. Tal conclusão foi similar à relatada por Reis *et al.* (2009), de que o nível “Com adoçante” causou impacto negativo em relação à “Sem adição de

açúcar”, em um estudo de *Conjoint Analysis* com avaliação das embalagens de iogurte baseada em notas.

Similarmente, a amostra de iogurte com “0% de gordura” foi 2,2000 vezes mais provável de ser escolhida do que a amostra “Com baixo teor de gordura”. Portanto, os consumidores entrevistados desejam consumir um produto isento de gordura.

A amostra de iogurte “Com Proteínas Bioativas” foi 4,5385 vezes mais provável de ser escolhida do que a amostra “Com Proteínas do soro do leite”, sugerindo que o consumidor associou um aspecto negativo ao termo “proteínas do soro do leite”, possivelmente devido às características sensoriais percebidas como pouco agradáveis. Por outro lado, conforme relatado por Della Lucial *et al.* (2010) o termo “bioativas” parece ter sido associado a um alimento mais saudável.

No enfoque Bayesiano, além das estimativas pontuais de Razão de Escolha, o desvio padrão a *posteriori* também foi obtido, o que permitiu avaliar a precisão dessas estimativas. As densidades estimadas das razões de escolha podem ser verificadas em ANEXO III. Pela análise dos intervalos H.P.D., novamente verificou-se diferença estatística entre o efeito dos níveis para os três atributos avaliados, uma vez que o valor unitário (1) não pertenceu aos respectivos intervalos. Adicionalmente, este tipo de análise pode ser útil nos estudos em que os atributos possuem mais de dois níveis, pois na CBCA, com o tradicional teste de Wald, infere-se apenas sobre a significância estatística dos atributos avaliados. Portanto, além de identificar quais os níveis exerceriam a maior e a menor influência na escolha do consumidor, seria possível realizar comparações estatísticas entre estes, comparando as razões de escolhas dentro de cada atributo.

4 Conclusões

Neste trabalho discutiu-se o enfoque Bayesiano para o modelo Logit Multinomial, empregado na CBCA. A análise proposta apresentou resultados similares aos do enfoque Frequentista, no entanto, permitiu acessar a distribuição a *posteriori* das probabilidades e razões de escolha e, conseqüentemente, testar hipóteses e realizar comparações estatísticas sobre as mesmas. Essa foi a principal vantagem do enfoque Bayesiano em relação ao enfoque Frequentista, no qual apenas as estimativas pontuais foram obtidas. Destaca-se ainda que o procedimento Bayesiano não exigiu ainda nenhum pressuposto sobre tamanho amostral e a maximização de uma função específica.

Em estudos futuros sugere-se a aplicação desse enfoque com distribuições de probabilidades *a priori* normais e informativas, para verificar se com o conhecimento prévio do pesquisador, os resultados poderiam ser distintos dos obtidos sob o enfoque Frequentista, especialmente com número reduzido de consumidores. Uma segunda sugestão seria verificar os resultados com a utilização de outros modelos de probabilidades *a priori* e com a função de verossimilhança original do modelo Logit Multinomial (sem a aproximação de Laplace).

Quanto aos aspectos práticos do estudo, concluiu-se que os três atributos manipulados na análise (açúcar, gordura e proteína) foram importantes na escolha dos consumidores. A amostra de iogurte *light* sabor morango contendo as informações “0% de açúcar”, “0% de gordura” e “enriquecido com proteínas bioativas” foi a que obteve maior probabilidade de escolha para o grupo de consumidores participantes deste estudo.

Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio financeiro. Adicionalmente, aos revisores anônimos e editores pelas valiosas sugestões.

BARBOSA, E. C.; SILVA, C. H. O.; NASCIMENTO, M.; SILVA, F. F.; MINIM, V. P. R.; DELIZA, R.; DELLA LUCIA, S. M. *Choice-Based Conjoint Analysis: A Bayesian approach*. *Rev. Bras. Biom.* Lavras, v.36, n.1, p.1-19, 2018.

- **ABSTRACT:** *In this paper we present a Bayesian approach to Frequentist Logit Multinomial model used in Choice-Based Conjoint Analysis. The analysis was conducted with data of choice for the assessment by 144 consumers, eight samples of light strawberry flavored yogurts, ranging information from three ingredients (sugar, fat and protein) in a full factorial design. The results and inferences obtained by the Bayesian approach are presented in terms of estimating the main effect of the attributes, the choice probabilities and choice ratio. The Frequentist results are also reported and discussed. The Bayesian analysis showed similar results to Frequentist and allowed the construction of credibility intervals for choice probabilities and choice ratio, allowing statistically compare such quantities. About the practical results, the most likely choice was associated with yoghurt containing strawberry flavor light information "0% sugar", "0%fat" and "bioactive proteins enriched".*
- **KEYWORDS:** *Multinomial Logit model; factorial; choice of yogurts.*

Referências

AAKER, D. A.; KUMAR, V.; DAY, G. S. *Marketing research*. John Wiley & Sons, 2008.

AIZAKI, H.; NISHIMURA, K. Design and analysis of choice experiments using R: a brief introduction. *Agricultural Information Research*, v.17, n.2, p.86-94, 2008.

ALLENBY, G.; ROSSI, P. Marketing models of consumer heterogeneity. *Journal of Econometrics*, v.89, p.57–78, 1999.

ANDERSEN, P. K.; GILL, R. D. Cox's regression model for counting processes: a large sample study. *The annals of statistics*, p.1100-1120, 1982.

ARENOE, B.; VAN DER REST, J. P. I.; KATTUMAN, P. Game theoretic pricing models in hotel revenue management: An equilibrium choice-based conjoint analysis approach. *Tourism Management*, v.51, p.96-102, 2015.

BARRETO, G.; ANDRADE, M. G. Robust Bayesian approach for AR(p) models applied to stream flow forecasting. *Journal of Applied Statistical Science*, v.12, p.269-292, 2004.

BAYES, P. M. An Essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFRS. *Philosophical Transactions*, p.370-418, 1763.

CHEN, M. H.; SHAO, Q. M. Monte Carlo estimation of Bayesian credible and H.P.D. intervals. *Journal of Computational and Graphical Statistics*, v.8, n.1, p.69-92, 1999.

- DAZIANO, R. A. Conditional-logit Bayes estimators for consumer valuation of electric vehicle driving range. *Resource and Energy Economics*, v.35, n.3, p.429-450, 2013.
- DELIZA, R.; MACFIE, H. J. The generation of sensory expectation by external cues and its effect on sensory perception and hedonic ratings: a review. *Journal of Sensory Studies*, v.11, n.2, p.103-128, 1996.
- DELIZA, R.; ROSENTHAL, A.; HEDDERLEY, D.; JAEGER, S. R. Consumer perception of irradiated fruit: a case study using choice-based conjoint analysis. *Journal of Sensory Studies*, v.25, n.2, p.184-200, 2010.
- DELIZA, R.; ROSENHAL, A.; SILVA, A. L. S. Consumer attitude towards information on non-conventional technology. *Trends in Food Science & Technology*, v.14, n.1-2, p.43-49, 2003.
- DELLA LUCIA, S. M.; MINIM, V. P. R.; SILVA, C. H. O.; MINIM, L. A.; SILVA, R. C. S. N. Análise conjunta de fatores baseada em escolhas no estudo da embalagem de iogurte light sabor morango. *Brazilian Journal of Food technology*, p.11-18, 2010.
- DOOB, J. L. Application of the theory of martingales. *Colloquy International Centre Nat. Rech. Sci*, p.22-28, 1949.
- EFRON, B. *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia, 1982.
- GALLANT, A. R. *Nonlinear statistical models*. New York: John Wiley & Sons, 2009.
- GAMERMAN, D.; LOPES, H. F. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. CRC Press, 2006.
- GEWEKE, J. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica: Journal of the Econometric Society*, p.1317-1339, 1989.
- GREEN, P. E.; KRIEGER, A. M. Conjoint analysis with product-positioning applications. *Handbooks in operations research and management science*, v.5, p.467-515, 1993.
- GREENE, W. H. *Econometric analysis*. Pearson Education India, 2003.
- KOOP, G.; POIRIER, D. J. Bayesian analysis of logit models using natural conjugate priors. *Journal of Econometrics*, v.56, n.3, p.323-340, 1993.
- LAPLACE, P. S. Memoir on the probability of the causes of events. *Statistical Science*, p.364-378, 1986.
- LECAM, L. Les propriétés asymptotiques des solutions de Bayes. *Publ. Inst. Statist. Univ. Paris*, v.7, p.18-35, 1958.
- LOUVIERE, J. J.; WOODWORTH, G. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of marketing research*, p.350-367, 1983.
- MACFIE, H. J.; BRATCHELL, N.; GREENHOFF, K.; VALLIS, L. V. 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, p.129-148, 1989.

MAHAJAN, V.; GREEN, P. E.; GOLDBERG, S. M. A conjoint model for measuring self-and cross-price/demand relationships. *Journal of Marketing Research*, p.334-342, 1982.

MCFADDEN, D Conditional Logit Analysis of Qualitative Choice Behavior, In: P. ZAREMBKA, P. (Ed.). *Frontiers in Econometrics*, New York: Academic Press, 1974, p. 105-142.

MEYER, R. J.; KAHN, B. E. Probabilistic models of consumer choice behavior. *Handbook of Consumer Behavior*, p.85-88, 1991.

MOORE, W. L. A cross-validity comparison of rating-based and *Choice-Based Conjoint Analysis* models. *International Journal of Research in Marketing*, v.21, n.3, p.299-312, 2004.

NATTER, M.; FEURSTEIN, M. Real world performance of choice-based conjoint models. *European Journal of Operational Research*, v.137, n.2, p.448-458, 2002.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria 2017. Disponível em: <http://www.r-project.org>.

RAFTERY, A. E.; LEWIS, S. How many iterations in the Gibbs sampler. *Bayesian statistics*, v.4, n.2, p.763-773, 1992.

REIS, R. C.; MINIM, V. P. R.; DIAS, B. R. P.; CHAVES, J. B. P.; MINIM, L. A. Impact of the use of different sweeteners in the acceptability of strawberry light yogurt. *Alimentos e Nutrição*, v.20, n.1, p.53-60, 2009.

ROSSI, P. E.; ALLENBY, G. M.; MCCULLOCH, R. *Bayesian statistics and marketing*. New York: John Wiley & Sons, 2005.

ROSSI, P.; MCCULLOCH, R.; ALLENBY, G The value of household information in target marketing. *Marketing Science*, v.15, p.321-340, 1996.

SO, Y.; KUHFIELD, W. F. *Multinomial logit models*. SUGI 20 Conference Proceedings. 1995.

SORENSEN, D. A.; WANG, C. S.; JENSEN, J. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genetics Selection Evolution*, v.26, n.4, p.333-360, 1994.

TEMPESTA, T.; GIANCRISTOFARO, R. A.; CORAIN, L.; SALMASO, L.; TOMASI, D.; BOATTO, V. The importance of landscape in wine quality perception: An integrated approach using choice-based conjoint analysis and combination-based permutation tests. *Food Quality and Preference*, v.21, n.7, p.827-836, 2010.

TIERNEY, L. Markov chains for exploring posterior distributions. *The Annals of Statistics*, p.1701-1728, 1994.

TRAIN, K. E. *Discrete choice methods with simulation*. Cambridge: Cambridge University Press, 2009.

UTZ, K. S.; HOOG, J.; WENTRUP, A.; BERG, S.; LÄMMER, A.; JAINSCH, B.; SCHENK, T. Patient preferences for disease-modifying drugs in multiple sclerosis

therapy: a choice-based conjoint analysis. *Therapeutic advances in neurological disorders*, v.7, n.6, p.263-275, 2014.

WALD, A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, v.54, p.426-482, 1943.

YUAN, A.; DE GOOIJER, J. G. Asymptotically informative prior for Bayesian analysis. *Communications in Statistics: Theory and Methods*, v.43, n.14, p.3080-3094, 2014.

ZELLNER, A.; ROSSI, P. E. Bayesian analysis of dichotomous quantal response models. *Journal of Econometrics*, v.25, p.365-394, 1984.

ZIMMERMANN, T. M.; CLOUTH, J.; ELOSGE, M.; HEURICH, M.; SCHNEIDER, E.; WILHELM, S.; WOLFRATH, A. Patient preferences for outcomes of depression treatment in Germany: a choice-based conjoint analysis study. *Journal of Affective Disorders*, v.148, n.2, p.210-219, 2013.

Recebido em 05.04.2016

Aprovado após revisão em 06.06.2017

ANEXO I

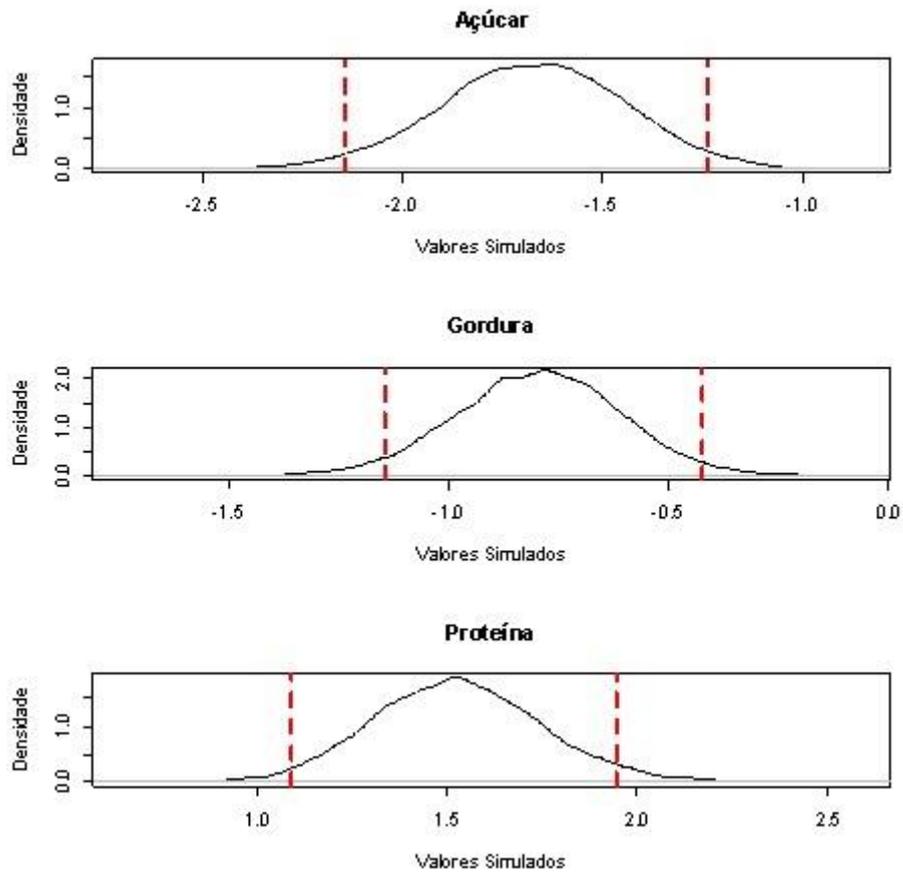


Figura 2 - Densidades a *posteriori* estimadas para os parâmetros (atributos) e os respectivos intervalos H. P. D. (linhas tracejadas em vermelho).

ANEXO II

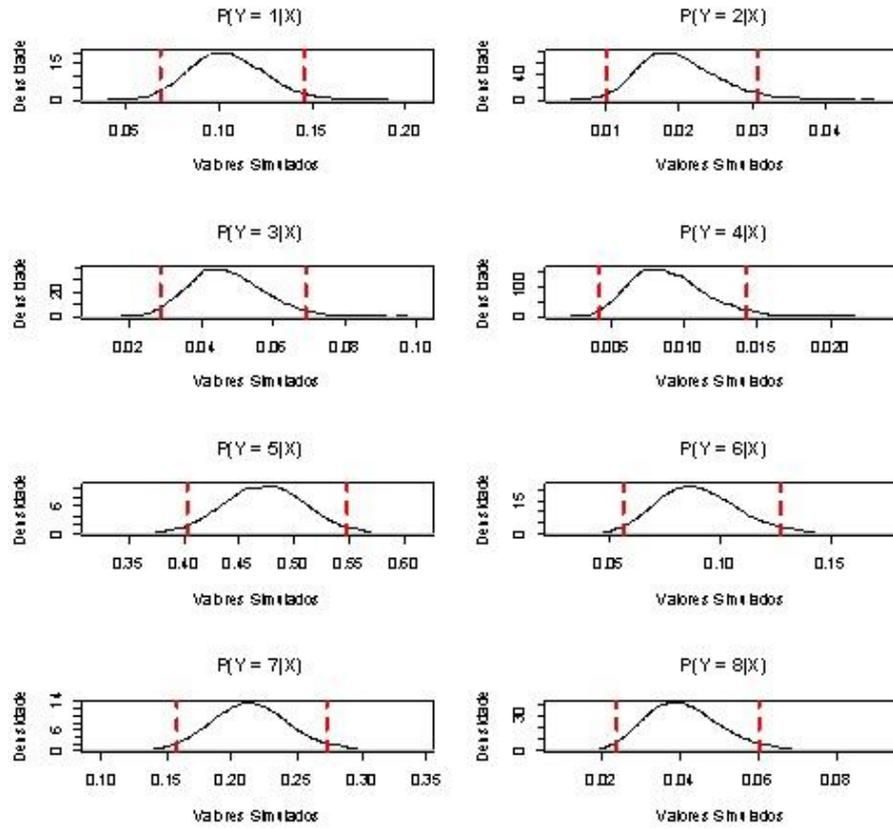


Figura 3 - Densidades *a posteriori* estimadas para as probabilidades de escolha e os respectivos intervalos H. P. D. (linhas tracejadas em vermelho).

ANEXO III

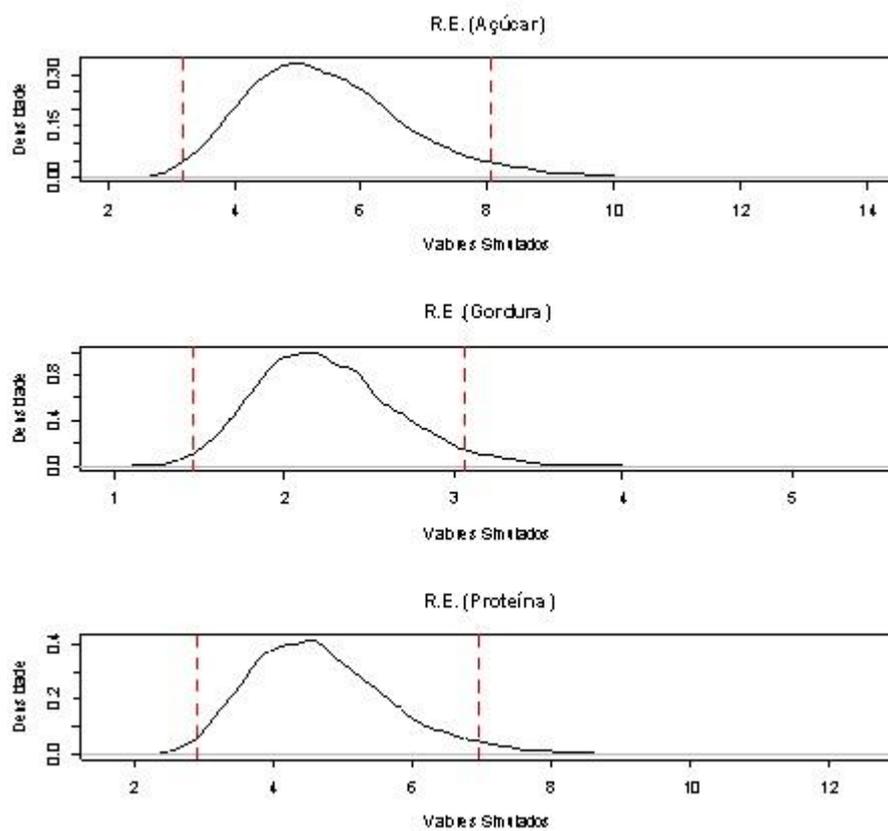


Figura 4 - Densidades *a posteriori* estimadas para as razões de escolha e os respectivos intervalos H. P. D. (linhas tracejadas em vermelho).