

# Abordagem bayesiana para análise de testes triangulares

Camilla Marques Barroso <sup>1 3</sup>

Júlio Sílvio de Sousa Bueno Filho <sup>2</sup>

## 1 Introdução

Testes de diferenciação podem ser usados para detectar fraudes em produtos alimentícios. Dentre estes testes destacam-se os testes triangulares como os mais utilizados por sua simplicidade de interpretação pelos pesquisadores, delineamento e implementação. Neste tipo de teste cada provador recebe três amostras codificadas, sendo duas iguais e uma diferente, devendo identificar a amostra diferente.

A análise usual considera que as proporções de respostas corretas seguem distribuição aproximadamente normal. No entanto, parece-nos mais adequado modelar diretamente o número de acertos usando a distribuição binomial em um contexto bayesiano.

Objetivou-se com este trabalho desenvolver uma abordagem bayesiana para o modelo usual de análise em testes triangulares na análise sensorial.

## 2 Material e métodos

Os dados utilizados nesse trabalho são provenientes de um experimento conduzido no esquema fatorial para avaliar a adição de diferentes tipos de adulterantes em amostras de café. Os fatores estudados foram: tipos de adulterantes em três níveis (casca de café, milho e palha de café) e proporção de adulteração em seis níveis (1%, 10%, 20%, 30%, 40% e 50%). O delineamento experimental foi em blocos casualizados, sendo que cada um dos vinte provadores constituiu um bloco (TAVARES, 2012).

Cada provador teve três chances para cada um dos dezoito tratamentos no teste triangular. Dessa forma, obteve-se a proporção de acertos para cada combinação de fatores.

A análise usual considera que essas proporções seguem distribuição aproximadamente normal. No entanto, parece-nos mais adequado modelar diretamente o número de acertos usando a distribuição binomial. Estudos mostram que o ajuste considerando este modelo é mais preciso e resulta em intervalos de confiança mais estreitos para os limiares de detecção (BARROSO, 2013).

---

<sup>1</sup>DEX - UFLA. e-mail: *camillamarques84@gmail.com*

<sup>3</sup>Agradecimento à CAPES e FAPEMIG pelo apoio financeiro.

<sup>2</sup>DEX - UFLA. e-mail: *jssbueno@dex.ufla.br*

O modelo binomial supõe que os provadores têm a mesma probabilidade de resposta, porém, essa suposição está em conflito com as ideias psicológicas sobre os processos de percepção e tomada de decisão. A estatística bayesiana, que trata um parâmetro como uma variável aleatória, em vez de uma constante desconhecida, parece fornecer a flexibilidade necessária para resolver esse problema. Esta é uma razão para se defender a abordagem bayesiana para testes sensoriais (BI, 2003).

Dessa forma, consideramos os blocos ou provadores como variáveis aleatórias e temos que a função de verossimilhança dos  $n$  ensaios referente a um experimento com  $b$  provadores para avaliar  $\tau$  tratamentos pode ser representada por:

$$P(y_{ij}/\pi_{ij}) = \prod_{i=1}^{\tau} \prod_{j=1}^b \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{m_{ij} - y_{ij}}$$

em que

$y_{ij}$ : número de respostas corretas do provador  $i$  no tratamento  $j$ ;

$m_{ij}$ : número de tentativas do provador  $i$  no tratamento  $j$ ;

$$\pi_{ij} = \frac{1}{3} + \frac{2}{3} \left( \frac{1}{1 + e^{-\eta_{ij}}} \right);$$

$$\eta = X\boldsymbol{\tau} + Z\mathbf{b};$$

$$\boldsymbol{\theta} = \left[ \boldsymbol{\tau}', \mathbf{b}' \right]';$$

$\boldsymbol{\tau}$  = vetor dos efeitos de tratamentos;

$\mathbf{b}$  = vetor dos efeitos de blocos;

As prioris estabelecidas para os parâmetros  $\boldsymbol{\tau}$ ,  $\mathbf{b}$  e  $\sigma_b^2$  são:

- $P(\boldsymbol{\tau}) \propto k$
- $\mathbf{b} \sim MVN(\boldsymbol{\theta}, I\sigma_b^2)$
- $\sigma_b^2 \sim \chi^{-2}(v, s_b^2)$

Supondo independência entre as distribuições *a priori* encontra-se a distribuição conjunta dos parâmetros *a posteriori* multiplicando-se a função de verossimilhança e as distribuições *a priori*.

$$P(\boldsymbol{\tau}, \mathbf{b}, \sigma_b^2, \eta | Y) \propto \left[ \prod_{i=1}^{\tau} \prod_{j=1}^b \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{m_{ij} - y_{ij}} \right] P(\mathbf{b}) P(\sigma_b^2)$$

A inferência sobre todos os parâmetros foram feitas usando aproximações das distribuições *a posteriori* conjuntas por meio de métodos de Monte Carlo via cadeias de Markov (MCMC). Esses métodos envolvem a simulação de cadeias de Markov construídas adequadamente que têm como distribuição de equilíbrio a distribuição de interesse (usualmente a posteriori conjunta), diretamente da qual não é possível gerar uma amostra.

A ideia por trás da simulação MCMC é construir uma densidade de transição que converge para a densidade posteriori a partir de qualquer ponto de partida  $\theta_0$ . As saídas obtidas são amostras correlacionadas da distribuição posteriori.

O algoritmo Metropolis-Hastings e o amostrador de Gibbs são exemplos desse método, no qual é necessário e suficiente que se conheça as distribuições condicionais completas.

O amostrador de Gibbs seleciona amostras aleatórias de todas as distribuições condicionais completas obtidas da distribuição *a posteriori* conjunta. Quando estas distribuições não são identificadas ou se a geração não iterativa for complicada usamos o algoritmo de Metropolis-Hastings.

A partir da distribuição conjunta *a posteriori*, obtém-se as distribuições condicionais completas de cada variável, fixando-se as demais variáveis:

- $P(\tau|b, \sigma_b^2, y) \propto \prod_{i=k} \prod_{j=1}^b \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{m_{ij} - y_{ij}}$
- $P(b|\tau, \sigma_b^2, y) \propto \prod_{j=k} \prod_{i=1}^r \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{m_{ij} - y_{ij}}$
- $\sigma_b^2|\tau, b, y \sim \chi^{-2} \left( q + v_b, \frac{b'b + v_b s_b^2}{q + v_b} \right)$

O critério utilizado para se verificar se a cadeia é estacionária foi o diagnóstico de convergência de Raftery-Lewis. Este teste indica o número de iterações iniciais que devem ser descartadas (“burn-in”) e a distância mínima de uma iteração à outra para se obter uma amostra aproximadamente independente (“jump”).

### 3 Resultados e discussão

No processo de geração das cadeias foi obtida uma amostra final de tamanho 4000 para cada parâmetro, após ter sido feito um descarte de 1000 valores iniciais e um salto de 150 iterações.

Pelo critério de Raftery e Lewis (1992) observa-se que os valores do fator de dependência foram menores que cinco para todos os parâmetros, condição suficiente para a garantia de convergência das cadeias. Além disso, os gráficos do traço para todos os parâmetros sugerem estacionariedade e convergência das cadeias.

O gráfico do traço e da densidade posteriori para o parâmetro  $\sigma_b^2$  estão apresentados na Figura 1. Observa-se pelo gráfico do traço que a cadeia se estabilizou com 4000 iterações e o gráfico da densidade posteriori do parâmetro  $\sigma_b^2$  mostra que a distribuição se aproxima na sua totalidade da distribuição normal.

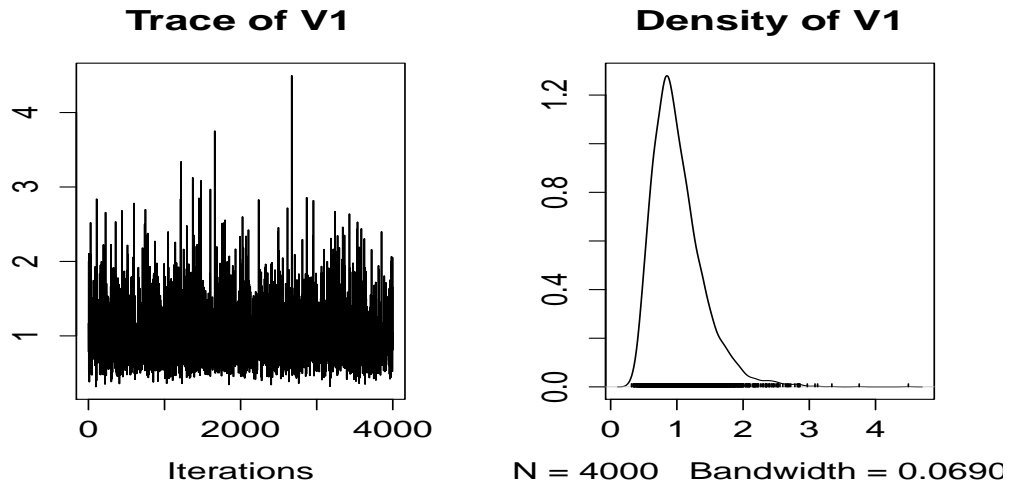


Figura 1: Representação gráfica do traço e da função densidade posteriori do parâmetro  $\sigma_b^2$

Encontrou-se o intervalo de credibilidade para cada parâmetro e apresentamos na Tabela 1 a estimativa da média e o intervalo HPD para o parâmetro  $\sigma_b^2$ .

Tabela 1: estimativa da média e o intervalo HPD para o parâmetro  $\sigma_b^2$

LI	Estimativa	LS
0,4127	1,0174	1,7805

Quanto maior esta variância maior a distinção entre os provedores em média. Quanto mais semelhantes forem as medidas de um provedor entre si, mais distintos serão os provedores em média.

A Figura 2, proporção de acertos em função da percentagem de adulteração, mostra que as curvas considerando os valores estimados, em vermelho, estão bem próximas das curvas considerando os valores observados, em preto. Obteve-se, assim, um bom ajuste para os três adulterantes.

Em trabalhos futuros faremos estudos de simulação para avaliar em quais aspectos as técnicas bayesianas possuem melhor desempenho.

## 4 Conclusões

Os resultados permitem concluir que o método bayesiano proposto apresentou bons resultados e pode ser utilizado na análise de testes triangulares em análise sensorial.

## Referências

- [1] BARROSO, C. M. **Proposta de análise estatística de testes triangulares considerando a distribuição Binomial**. 62 p. Dissertação (Mestrado em Estatística e Experimentação

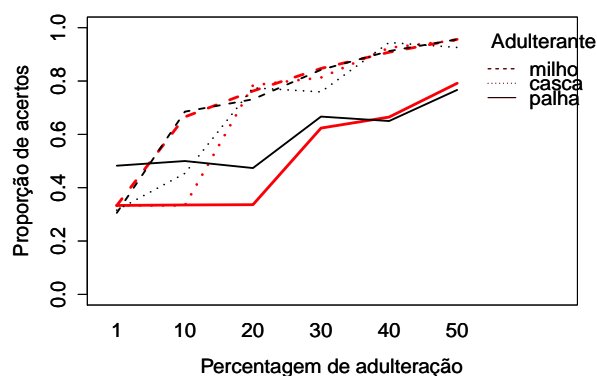


Figura 2: Representação gráfica da proporção de acertos observada e estimada em função da porcentagem de adulteração para cada tipo de adulterante

Agropecuária) Universidade Federal de Lavras, Lavras, 2013.

- [2] TAVARES, K. M. **Detecção de adulteração em café riado torrado e moído por espectroscopia no infravermelho e análise sensorial**. 2012. 109 p. Dissertação (Mestrado em Ciência dos Alimentos) Universidade Federal de Lavras, Lavras, 2012.
- [3] BI, J. Difficulties and a way out: a bayesian approach for sensory difference and preference tests. **Journal of Sensory Studies**. Richmond, v.18, n.1, p. 1-18, 2003.
- [4] RAFTERY, A.E. LEWIS, S.M. One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. **Statistical Science**, p. 493-497, 1992.