REVISTA
DO INSTITUTO DE LATICÍNIOS CÂNDIDO TOSTES
v. 77, n. 3, jul-set 2022

Artigo DOI: 10.14295/2238-6416.v77i3.892

ANÁLISE SENSORIAL DE QUEIJOS PRATO VIA MODELO LINEAR MISTO GENERALIZADO

Sensory analysis of Prato cheeses by generalized linear mixed models

Tatiane Carvalho Alvarenga[1], Jéssica Ferreira Rodrigues[2]*

RESUMO

A análise sensorial, área da Ciência dos Alimentos, é empregada para analisar e medir características dos alimentos, podendo avaliar quanto à aceitação das amostras. Tais avaliações podem ser realizadas por meio da escala hedônica numérica de 9 pontos, classificada como variável categorizada ordinal. Com o objetivo de verificar a significância de cada efeito envolvido na análise sensorial, foi realizado no Departamento de Ciências dos Alimentos da Universidade Federal de Lavras, um experimento com duas marcas de queijos Prato, avaliadas por 100 avaliadores de cada sexo masculino e feminino. Para isso foi utilizado modelos de efeito fixo e aleatório, considerando os efeitos de queijo e sexo e a interação entre ambos de efeitos fixos, e o efeito dos diferentes avaliadores considerado como efeito aleatório, pelo fato de serem as repetições no experimento. Concluiu-se que os efeitos de queijo Prato, sexo e a interação entre estes e o efeito do avaliador foram todos significativos no modelo (AIC=1465), mostrando que o melhor modelo para a análise das notas quanto à aceitação dos queijos, foi o modelo misto generalizado. A marca A teve melhor aceitação quanto a preferência (nota média na escala hedônica de 7 pontos).

Palavras-chave: análise de dados categorizados; modelos de chances proporcionais. modelos encaixados.

ABSTRACT

Sensory analysis, an area of Food Science, is used to analyze and measure characteristics of foods, being able to evaluate the acceptance of samples. Such assessments can be performed using the 9-point numerical hedonic scale, classified as an ordinal categorized variable. To verify the significance of each effect involved in the sensorial analysis, the experiment was carried out at the Department of Food Science at the Federal University of Lavras, with two brands of Prato cheese evaluated by 100 evaluators, each male and female. For this, fixed

1 Universidade Estadual de Maringá, Departamento de Estatística, Avenida Colombo, 5790, Zona 7, 87020-900, Maringá, PR, Brasil. E-mail: tatianecarvalhoalvarenga@gmail.com
2 Universidade Federal de Lavras, Departamento de Ciência dos Alimentos, Lavras, MG, Brasil.

*Autor para correspondência

Recebido / Received: 11/10/2021 ***Aprovado / Approved: 31/03/2022***

and random effect models were used, considering the effects of cheese and sex and the interaction between both fixed effects, and the effect of the different evaluators was considered as a random effect because they are the repetitions in the experiment. It was concluded that the effects of cheese, sex, and the interaction between them and the evaluator effect were all significant in the model (AIC=1465), showing that the best model for analyzing the grades regarding the acceptance of cheeses was the generalized mixed model. Brand A had better acceptance in terms of preference (average score on the 7-point hedonic scale).

Keywords: analysis of categorized data; proportional odds models; nested models.

INTRODUÇÃO

Os dados categorizados são encontrados nas diferentes áreas de conhecimento (PAULINO; SINGER, 2006), seja, nas Ciências Exatas e da Terra, Ciências Biológicas, Engenharias, Ciências da Saúde, Ciências Agrárias, Ciências Sociais Aplicadas, Ciências Humanas, Linguística, Letras e Artes, entre outras. Esse tipo de dado é classificado em nominal, quando sua variável resposta atribui uma característica à amostra, não passando a ideia de ordenação, e ordinal se existe uma suposição de ordem. Nas Ciências dos Alimentos é comum esse tipo de variável, como por exemplo na análise sensorial: avaliação de um determinado sabor de acordo com a escala hedônica numérica de 9 pontos (IFT, 1981) (variável ordinal), ou característica sensorial de um produto (variável nominal).

As informações contidas nos experimentos envolvendo os dados categóricos podem ser expressas por modelos estatísticos. Esses são classificados de acordo com efeito envolvido, seja fixo, aleatório ou misto (SEARLE, 1992). Se o efeito é fixo o modelo é fixo, caso o efeito envolvido seja aleatório o modelo é aleatório, e quando ambos os feitos então envolvidos no modelo ele é classificado como misto (SEARLE, 1992).

De acordo com Barbin (1993), o efeito é fixo quando as considerações sobre os tratamentos são limitadas a amostra selecionada, que tem por finalidade estimar os efeitos dos tratamentos. Ao contrário, é visto quando o efeito é aleatório, ou seja, os tratamentos selecionados de uma amostra representam uma população, na qual tem por finalidade predizer as componentes de variância. Os modelos mistos são aqueles que além de efeitos fixos; ao menos o erro aleatório associado às observações do experimento; possui efeitos aleatórios, na qual se pode atribuir uma distribuição de probabilidade a eles.

O modelo estatístico sempre tem a componente da média (μ) que é sempre fixa e a componente do erro (e_{ij}) que é sempre aleatória, mas ambos os efeitos não entram na classificação do modelo, apenas os outros efeitos (tratamentos, bloco, interação, entre outros) (McCULLOCH; SEARLE, 2001).

Quando se trabalha com variáveis ordinais, os modelos lineares clássicos não são os mais indicados, assim, opta-se por uma generalização desses, nomeados de modelos lineares generalizados, devido a variável resposta ser categorizada e os modelos generalizados podem ter suas variáveis respostas em qualquer família exponencial de distribuição (modelos log-lineares, modelos gama para dados contínuos, modelos logito e probito para dados de proporções, dentre outros). Estes modelos são definidos por três componentes: a distribuição da variável resposta, o preditor linear e a função de ligação. O preditor linear é a combinação das variáveis independentes com os parâmetros $\left(\eta = x^{'}\beta\right)$. Já a função de ligação vincula o preditor linear a média populacional $\left(g(\mu_i) = \eta_i = x_i^{'}\beta\right)$ (McCULLOCH; SEARLE, 2001).

Modelos de regressão para dados categorizados ordinais

Na literatura os modelos utilizados para variáveis respostas ordinais, são uma extensão dos Modelos Lineares generalizados para o caso multivariado, atuando como componente sistemático do modelo. Os mais comuns seguem descritos a seguir.

Modelos de logitos cumulativos

Uma característica importante quando se trabalha com dados ordinais é a possibilidade de utilizar as probabilidades acumulativas, na qual se utiliza logitos de probabilidades de respostas acumuladas, denominados de logitos cumulativos, definidos como:

$$Logito\ [P(Y \leq j|x)] = log\left[\frac{P(Y \leq j|x)}{1-P(Y \leq j|x)}\right],$$

em que $P(Y \leq j|x)$ é a probabilidade acumulada de Y, ou seja, toda a probabilidade inferior a *j-ésima* categoria.

Quando o número de categorias é maior do que 2, $J > 2$, exitem $J-1$ logitos, podendo assim, ser expresso pelo modelo:

$$Logito\ [P(Y \leq j|x)] = \alpha_j + \sum_{k=1}^{p} \beta_{jk} x_k$$
$$= \alpha_j + \beta_j^{'} x,$$
$$j = 1,...J-1,$$

em que *p* é o número de covariáveis, $\beta_j = (\beta_{j1},..., \beta_{jp})^{'}$ é o vetor de parâmetros e $x = (x_1,..., x_p)^{'}$ representa o conjunto de valores das *p*-covariáveis do modelo.

Modelos de chances proporcionais

McCullagh (1980) propôs os Modelos de chances proporcionais, na qual as proporções são iguais, ou seja, $\beta_j = \beta$ para todo j. Sendo definido por:

$$Logito\ [P(Y \leq j|x)] = \alpha_j + \sum_{k=1}^{p} \beta_k x_k$$
$$= \alpha_j + \beta^{'}\ x,$$
$$j = 1,...,J-1,$$

sendo os efeitos de x idênticos para todos os logitos cumulativos.

Este trabalho tem como objetivos: 1) Ajustar modelo misto, aleatório e fixo, para a avaliação de duas marcas de queijo Prato do Sul de Minas Gerais. 2) Selecionar o melhor modelo. 3) Verificar a significância de cada efeito ao modelo selecionado.

MATERIAL E MÉTODOS

Materiais

Por meio do teste de aceitação com a apresentação das amostras em ordem balanceada duas marcas de queijo Prato tradicional, A e B, foram analisadas por 100 provadores de cada sexo, sendo este experimento realizado no laboratório de Análise Sensorial, no Departamento de Ciência de Alimentos da Universidade Federal de Lavras. As marcas foram avaliadas por meio da escala hedônica numérica de 9 pontos (IFT, 1981), atribuindo notas entre 1 e 9, sendo a nota 9 a melhor avaliação quanto a preferência.

As análises estatísticas foram realizadas pelos pacotes *vcrpart* e ordinal do *software* R (R *Core Team*, 2021).

Métodos

Foram utilizados os modelos de regressão para dados categorizados ordinais e aplicou-se o teste da razão de verossimilhança para verificar a suposição de proporcionalidade, considerando H_0: $\beta_j = \beta$.

A partir do processo sequencial, obteve-se as covariáveis significativas no modelo.

RESULTADOS E DISCUSSÃO

Análise descritiva

Inicialmente foi realizada uma análise descritiva dos dados, ou seja, a pontuação dos 200 avaliadores com relação às duas marcas de queijo Prato.

Ajuste do modelo

Pelo teste da razão de verossimilhança a suposição de proporcionalidade foi atendida (valor- p = 0,09939). Os efeitos de queijo Prato, sexo e a interação entre estes foram considerados fixos e o efeito de avaliador foi considerado aleatório.

Para comparação do melhor modelo utilizou-se o Critério de Informação de Akaike (AIC). Apesar da proximidade destes valores pode-se observar na Tabela 2, que o modelo 1 produziu o menor valor de AIC, indicando um melhor ajuste. Assim, foi estimado o efeito de cada parâmetro, que é verificado na Tabela 3.

A interpretação dos parâmetros do modelo de regressão baseia-se nas relações entre as notas atribuídas pelos avaliadores.

Tabela 1. Análise descritiva das avaliações de cada marca de queijo Prato

Marcas / Notas	2	3	4	5	6	7	8	9	Total	Nota média
A	6	6	12	7	32	32	68	37	200	7,0
B	17	30	35	11	20	36	46	5	200	5,5

Tabela 2. Modelos para experimento em estudo

Modelo	Número de parâmetros	AIC
1: Queijo Prato + Avaliador + Sexo + Sexo * Queijo Prato	11	1465,0
2: Queijo Prato + Avaliador + Sexo	10	1468,7
3: Avaliador + Sexo	9	1539,5
4: Queijo Prato + Avaliador	9	1477,7
5: Avaliador	8	1546,5

Tabela 3. Estimativas para os parâmetros do modelo 1

Parâmetro	Estimativa	Erro padrão	Valor z
Intercepto 2	-4,52746	0,52877	-8,5623
Intercepto 3	-3,27712	0,51288	-6,3896
Intercepto 4	-2,30018	0,48102	-4,7818
Intercepto 5	-1,99407	0,47964	-4,1574
Intercepto 6	-1,22659	0,45934	-2,6704
Intercepto 7	-0,27824	0,43547	-0,6389
Intercepto 8	1,91336	0,39535	4,8396

CONCLUSÃO

O melhor modelo que justificou o comportamento dos dados foi o que envolveu todos os efeitos (queijo Prato, avaliador, sexo e a interação entre sexo e queijo Prato), ou seja, o modelo misto. Os efeitos de queijo Prato, sexo e a interação entre estes foram todos significativos. A marca A teve melhor aceitação quanto a preferência.

AGRADECIMENTOS

À CAPES (Comissão de Aperfeiçoamento de Pessoal do Nível Superior) e FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais), pelo apoio financeiro.

REFERÊNCIAS

BARBIN, D. **Componentes de Variância**: Teoria e Aplicações. Piracicaba: FEALQ. 1993. 120 p.

IFT - Institute of Food Technologists. Sensory evaluation guide for testing food and beverage products. **Food technology**, v. 35, p. 50-59, 1981.

McCULLAGH, P. Regression models for ordinal data. **Journal of the Royal Statistical Society**, Series B (Methodological), v. 42, p. 109-142, 1980.

McCULLOCH, C. E., SEARLE, S. R. **Generalized, Linear and Mixed Models**. Nova Iorque: Willey. 2001. 325p.

PAULINO, C. D. M.; SINGER, J. M. **Análise de Dados Categorizados**. São Paulo: Edgard Blucher, 2006. 648p.

R Core Team. **R**: A language and environment for statistical computing. Viena: **R Foundation for Statistical Computing**, 2021. Disponível em: http://www.R-project.org

SEARLE, S. R.; CASELA, G.; McCULLOCH, C. E. **Variance Component**. Nova Iorque: John Wiley & Sons, 1992. 501 p.