

Biplots corrigidos por predições do modelo GEE-Logit para estudo de efeitos de safras na classificação granulométrica de cafés

**Haiany Aparecida Ferreira¹, Matheus Feres Freitas², Carla Regina Guimarães Brighenti³,
Marcelo Ângelo Cirillo⁴, Mírian Rosa⁵**

1. Introdução

A importância da comercialização e produtividade do café brasileiro vem ganhando espaço no setor econômico, graças ao volume de exportação, o qual, dado um destaque, a dinâmica dos preços internacionais, associado ao forte crescimento do mercado consumidor. Diante disso, o incentivo a produção e a classificação do produto com o intuito de obter uma bebida de qualidade mais elevada, considerando o café beneficiado têm crescido.

A classificação é feita com base na avaliação da qualidade e aspecto físico dos grãos de café, para isso geralmente utiliza-se a granulometria, mencionada por Brighenti e Cirillo (2018) e Brasil (2003) essa classificação engloba três etapas: tipos e defeitos, características da qualidade do grão e a qualidade da bebida. Os tipos e defeitos consistem em analisar a coloração e a contagem dos grãos defeituosos (defeitos intrínsecos) ou das impurezas (defeitos extrínsecos), seguindo a tabela de Classificação Oficial Brasileira (COB), em que cada amostra consiste em 300g de grãos beneficiados.

Os defeitos intrínsecos estão relacionados à aplicação inadequada das práticas de produção e industrial e modificações de origem genética ou fisiológica. São considerados defeitos intrínsecos: grãos ardidos, pretos, conchas, verdes, brocados e mal granados ou chochos. Quanto aos defeitos extrínsecos ou impurezas, são ocasionados por elementos que não pertencem ao café, dados por fatores como: secagem excessiva, mistura de café do chão com o café do pano, máquinas de colheita mal regulada, dentre outros. São considerados defeitos extrínsecos: cascas, paus, torrões, pedras, dentre outros (BRIGHENTTI; CIRILLO, 2018; CUSTODIO et al., 2007).

Segundo o Manual de Classificação do Centro de Comercio de Café do Estado de Minas Gerais (CCCMG, 2018) o número de grãos defeituosos avaliados seguindo a tabela de Classificação Oficial Brasileira (COB) é contado a partir de cada amostra e irá determinar o tipo de café. Para isso, a apuração do número de grãos defeituosos é feita em amostras de 300 gramas cada e estes são convertidos em defeitos conforme a tabela de equivalência. Esse contagem de defeitos tem seu resultado comparado a uma escala de sete níveis, em que cada nível corresponde a um tipo de café. A indicação é numérica e varia desde o tipo 2, café menos defeituoso, até o tipo 8, café mais defeituoso (SANTOS; NANTES, 2014).

Em consonância com os resultados de contagens dos grãos em diferentes categorias, obtidos na classificação dos defeitos, requer uma estrutura de dados, organizadas em tabelas de contingências. Uma possível organização das contagens dos defeitos classificados por safras e

¹Doutoranda em Estatística e Experimentação Agropecuária no DES/UFLA. email: haianyferreira@yahoo.com.br.

²Doutorando em Estatística e Experimentação Agropecuária no DES/UFLA. email: matheus712@hotmail.com.

³Professora Associada II, DZOO/UFSJ. email: carlabrighenti@ufs.br.

⁴Professor Associado III, DES/UFLA. email: macufla@dex.ufla.br.

⁵Mestranda em Estatística e Experimentação Agropecuária no DES/UFLA. email: mirianrosad@gmail.com.

porcentagens dos grãos de peneiras em nessas tabelas, tem em sua estrutura a caracterização de um delineamento em blocos com respostas multivariadas. Contudo, devido às técnicas embasadas na análise de variância serem infactíveis a natureza das respostas, sugere uma modelagem mais apurada. Diante disso, o objetivo deste trabalho é propor o uso de modelos de equações de estimação generalizada (LIANG; ZEGGER, 1986) em conjunto com a técnica biplot (GREENACRE, 2003) como um estudo complementar a análise granulométrica de grãos de cafés.

2. Material e Métodos

2.1 Material

O conjunto de dados utilizado foi encontrado em Brighenti e Cirillo (2018) e se refere a 40 amostras de cafés beneficiados da variedade catuaí. As amostras foram provenientes de colheitas feitas por produtores rurais do Sul de Minas nos anos de 2014 e 2015.

Segundo os autores, a classificação de defeitos nos grãos de café de cada amostra seguiu a tabela de Classificação Oficial Brasileira (COB) considerando a proporção dos grãos defeituosos ou das impurezas contidas em uma amostra de 300g de café. A classificação das amostras foi organizada em cinco defeitos codificados d1 a d5, cuja descrição é dada na tabela 1, com tripla entrada, dada as contagens, em relação aos tipos de defeitos e porcentagens dos grãos em peneiras, considerando as amostras coletadas nas safras de 2014 e 2015.

Tabela 1 - Contagens de defeitos em função da porcentagem (p) de grãos de peneira (17/18) referentes as safras de 2014 e 2015.

Safr	Tipo de Defeito	Porcentagem (p) de grãos em peneiras (17/18)		
		p<20%	20% ≤ p ≤ 30%	p>30%
2014	d1 (brocados)	187	223	151
	d2 (pretos, verdes e ardidos)	744	916	759
	d3 (quebrados)	588	498	465
	d4 (conchas)	127	252	156
	d5 (impureza)	12	8	12
2015	d1 (brocados)	264	190	274
	d2 (pretos, verdes e ardidos)	474	552	725
	d3 (quebrados)	324	430	262
	d4 (conchas)	210	293	382
	d5 (impureza)	12	0	53

Fonte: Adaptado de Brighenti e Cirillo (2018).

2.2 Métodos

Seguindo a estrutura dos dados apresentados anteriormente na Tabela 1. Para incorporar o efeito de safra associado aos defeitos foi utilizada a técnica biplot. Para isso inicialmente

considerou-se os valores preditos g_{ijk} obtidos através de 1000 realizações Monte Carlo no software R Core Team (2016), sendo esses organizados em uma tabela de tripla entrada, conseguidos pelo ajuste do modelo binomial com função de ligação logit de Equações de Estimação Generalizadas (*Generalized Estimating Equations* - GEE). Dado pela solução do sistema (1):

$$\sum_{i=1}^N \left(\frac{\partial \mu_i(\beta)}{\partial \beta} \right) V_i^{-1} (Y_i - \hat{\mu}_i(\beta)) = 0 \quad (1)$$

em que, $\hat{\mu}_i(\beta)$ correspondeu ao vetor de médias ajustados considerando os modelos binomiais com função de ligação logit (2):

$$\log \left(\frac{\mu_{ijk}}{1 - \mu_{ijk}} \right) = \eta_{ijk} ; \quad \mu_{ijk} = \frac{\exp(\eta_{ijk})}{1 + \exp(\eta_{ijk})} \quad (2)$$

Para a construção dos biplots as coordenadas foram obtidas através da decomposição de valores singulares nas matrizes blocos combinadas, em que, cada bloco corresponde às avaliações granulométricas dadas as amostras coletadas nas safras de 2014 e 2015, nomeadas em S_{2014} e S_{2015} , feita conforme procedimento sugerido por Greenacre (2003).

Desse modo, as componentes referentes ao efeito de safras, foram determinadas por: $S_{2014}+S_{2015}$ e $S_{2014}-S_{2015}$, correspondendo ao efeito combinado e a diferença entre as referidas safras.

3. Resultados e Discussão

Os biplots centrados nos valores preditos do modelo GEE com função de ligação logit, obtidos considerando o efeito de soma, $S_{2014}+S_{2015}$, e diferença, $S_{2014}-S_{2015}$, entre as safras foram obtidos conforme apresentado pela figura a seguir:

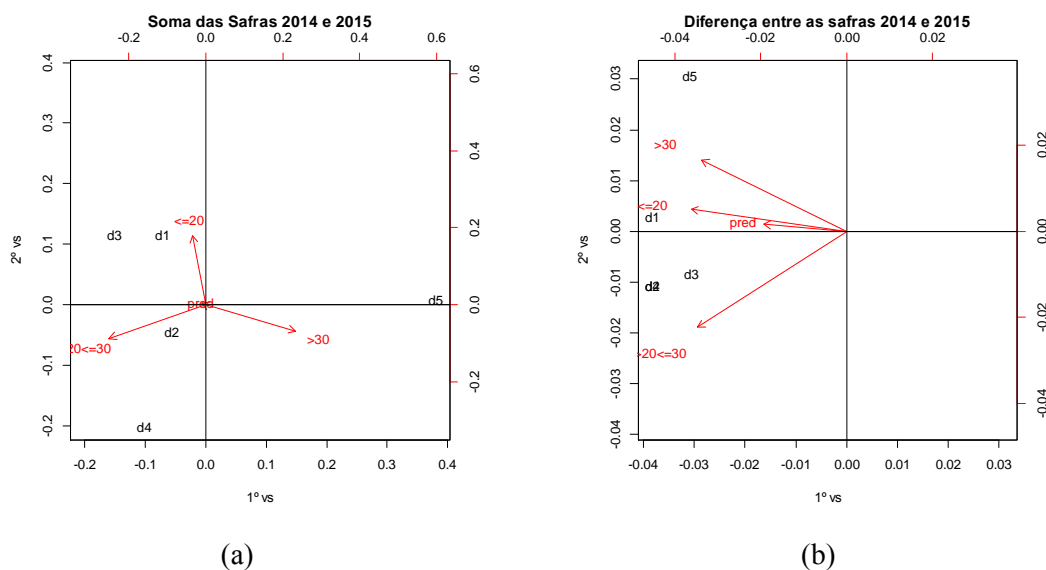


Figura 1 - Biplots centrados nos valores preditos do modelo GEE com função de ligação logit, $S_{2014}+S_{2015}$ (a) e $S_{2014}-S_{2015}$ (b).

Fonte: Do autor (2019).

Neste caso, a interpretação será focada no 1º e no 3º quadrante em ambos os biplots com o intuito de facilitar na hora da interpretação sugerindo uma discriminação entre um aspecto positivo ou negativo das associações. Mediante isto, observou-se que ao considerar o efeito aditivo das safras, $S_{2014} + S_{2015}$ (Figura 1 (a)), percebe-se que o efeito d2 (PVA) está caracterizado na classificação obtida a $20\% \leq p \leq 30\%$, o que em termos práticos indica uma maior incidência de grãos pretos, verdes e ardidos neste valor de porcentagem.

Em relação ao efeito da diferença das safras, $S_{2014} - S_{2015}$, o resultado da associação dos defeitos em relação a contagens em peneiras percebe-se a presença dos defeitos d2, d3 e d4 que correspondem respectivamente aos grãos PVA, quebrados e conchas que estão caracterizados em $20\% \leq p \leq 30\%$.

Dessa maneira tem-se que em ambos os casos observou-se a presença de grãos PVA em $20\% \leq p \leq 30\%$, o que condiz com a literatura existente, a citar o trabalho de Brighenti e Cirillo (2018) o qual utilizaram o modelo log-linear Hierárquico com estrutura de correlação independente e também verificaram que o defeito d2 (PVA) estava associado, as classificações de peneira $p < 20\%$ e $20\% \leq p \leq 30\%$.

4. Conclusões

Assim, com base nas observações obtidas quanto aos biplots correspondentes ao efeito aditivo das safras, $S_{2014} + S_{2015}$ e ao efeito da diferença, $S_{2014} - S_{2015}$ foi possível verificar em ambos os casos uma maior incidência de grãos pretos, verdes e ardidos com a porcentagem (p) de grãos em peneiras (17/18) caracterizada em $20\% \leq p \leq 30\%$, condizendo assim com a literatura existente.

Salientando ainda que os grãos defeituosos de cafés representam cerca de 15 a 20% da produção de café (Esquivel e Jiménez, 2012), os grãos pretos, verdes e ardidos (PVA) são considerados os piores defeitos, pois afetam diretamente a qualidade e o tipo de café. Contudo, tem sua porcentagem de defeitos equivalente ao período das colheitas, apresentando maior incidência de verdes no início e de pretos no final.

5. Agradecimentos

Agradecimentos à Cnpq e Fapemig - Edital Universal APQ -0242-16.

6. Referências Bibliográficas

- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa n 08, de 11 de junho de 2003. Disponível em: República Federativa do Brasil, Brasília, p.22-29, 20 ago 2003.
- BRIGHENTI, C. R. G.; CIRILLO, M. A. Analysis of defects in coffee beans compared to biplots for simultaneous tables. *Revista Ciência Agronômica*, v. 49, n. 1, p. 62-69, jan/mar, 2018.
- CENTRO DE COMERCIO DE CAFÉ DO ESTADO DE MINAS GERAIS. Minas Gerais. Manual de Classificação: Métodos de Classificação de Café utilizados pelo CCCMG, 2018.
- CUSTODIO, Anselmo A. de P.; GOMES, Natalino M.; LIMA, Luiz A.. Efeito da irrigação sobre a classificação do café. *Eng. Agríc., Jaboticabal*, v. 27, n. 3, p. 391-701, Dec. 2007.
- ESQUIVEL, P.; JIMÉNEZ, V. M. Functional properties of coffee and coffee by products. *Food Research International*, v. 46, n. 2, p. 488–495, Maio 2012.
- GREENACRE, M. J. Singular value decomposition of matched matrices. *Journal of Applied Statistics*, v. 30, n.10, 1101-1113. 2003.
- LIANG, K. Y.; ZEGGER, S. L. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, v. 1, n. 73, p. 13-22. 1986
- R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria, 2016. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>>.
- SANTOS, F. L.; NANTES, J. F. D. Coordenação no mercado do café brasileiro: o desserviço da classificação por defeitos. *Gestão e Produção, São Carlos*, v. 21, n. 3, p. 586-599, 2014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2014000300011&lng=en&nrm=iso>. Acesso em: 04 nov. 2018. <http://dx.doi.org/10.1590/S0104-530X2014000300011>.