

Monte Carlo simulation and importance sampling applied to sensory analysis validation of specialty coffees¹

Simulação monte carlo e amostragem por importância aplicada à análise sensorial validada da qualidade de cafés especiais

Haiany Aparecida Ferreira², Gilberto Rodrigues Liska³, Marcelo Ângelo Cirillo^{2*}, Flávio Meira Borém⁴, Diego Egídio Ribeiro⁴ and Ricardo Miguel Cortez⁴

ABSTRACT - Coffee sensory analysis is usually made by a sensory panel, which is formed by trained tasters, following the recommendations of the Specialty Coffee Association of America. However, the preference for a coffee is commonly determined by experimentation with consumers, who typically have no special skills in terms of sensory characteristics. Therefore, this study aimed at applying an intensive computational method to study sensory notes given by an untrained sensory panel, considering the probability distributions of the class of extreme values. Four types of specialty coffees produced under different processes and in varied altitudes in the mountainous region of Mantiqueira, Minas Gerais, were considered. We concluded that the generalized Pareto distribution can be applied to sensory analysis to discriminate types of specialty coffees. Furthermore, the method of importance sampling by Monte Carlo simulation showed greater variability considering a probabilistic model adjusted to identify specialty coffees.

Key words: Extreme Value Theory. Serra da Mantiqueira. Altitude. Consumers.

RESUMO - A análise sensorial do Café supõe que um painel sensorial é formado por provadores treinados, de acordo com as recomendações da American Specialty Coffee Association. No entanto, a escolha determina que a preferência de um café é rotineiramente feita através da experimentação com os consumidores, que em grande parte não tem habilidades especiais em termos de características sensoriais. Por este fato, este estudo objetivou aplicar o método computacional intensivo no estudo de notas sensoriais a partir de um painel sensorial não treinado considerando as distribuições de probabilidade pertencentes à classe dos valores extremos. Assim, foram considerados quatro tipos de cafés especiais produzidos em diferentes processos e alturas na região serrana da Mantiqueira, em Minas Gerais. Concluiu-se que a distribuição generalizada de Pareto pode ser aplicada à análise sensorial para discriminar os diferentes tipos de cafés especiais e que o método de amostragem por importância por simulação de Monte Carlo, considerando o modelo probabilístico ajustado para identificar o café especial, com notas apresentaram a maior variabilidade.

Palavras-chave: Valores Extremos. Serra da Mantiqueira. Altitude. Consumidores.

DOI: 10.5935/1806-6690.20210020

Editor-in-Article: Prof. Daniel Albiero - daniel.albiero@gmail.com

*Author for correspondence

Received for publication 05/06/2019; approved on 08/03/2021

¹Pesquisa financiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico/CNPq. Processo: 305599/2018-6

²Departamento de Estatística, Universidade Federal de Lavras (UFLA), 37200-000, Lavras-MG, Brasil, haianyferreira@yahoo.com.br (ORCID ID 0000-0001-5474-4441), macufla @ dex.ufla.br (ORCID ID 0000-0002-2647-439X)

³Departamento de Tecnologia Agroindustrial e Socioeconomia Rural (DTAiSeR), Centro de Ciências Agrárias, Universidade Federal de São Carlos, São Paulo-SP, Brasil, gilbertoliska@ufscar.br (ORCID ID 0000-0002-5108-377X)

⁴Departamento de Engenharia, Universidade Federal de Lavras (UFLA), Lavras-MG, Brasil, flavioiborem@deg.ufla.br (ORCID ID 0000-0002-6560-8792), diegoagro@hotmail.com (ORCID ID 0000-0002-1145-8661), ricardocortez@hotmail.com

INTRODUCTION

According to Ramos *et al.* (2016), a coffee is considered specialty when it has superior quality compared to its competitors in terms of origin, absence of defects, processing, and/or sensory characteristics, such as aroma and flavor, which are exclusively or essentially influenced by geography, as well as natural and chemical factors (CHAGAS *et al.*, 2013; MALTA; CHAGAS, 2009). Therefore, given the number of variation sources, computer simulation techniques involving methods of statistical data analysis have been widely used in the field of sensory quality.

Regarding the sensory quality profile of specialty coffees, which is associated with genetic and environmental factors (BORÉM *et al.*, 2016; RIBEIRO *et al.*, 2016). Ramos *et al.* (2016), used the chi-squared automatic interaction detection (CHAID) method to construct decision trees based on classifiers that related the sensory attributes to the environmental characteristics of the Serra da Mantiqueira region. The authors observed that CHAID method provided promising results regarding accuracy and hit rates by discriminating samples of *Coffea arabica* coffees, whose sensory evaluations, with scores ≥ 88 points, were characterized by production at altitudes $\geq 1,200$ m, with body intensity discriminated into high and low.

To identify similarities among four specialty coffees, Ossani *et al.* (2017), used multiple factor analysis for contingency tables with categorized data obtained from a sensory experiment conducted with different consumer groups. Despite their heterogeneity, the consumers involved in the analysis were successful at discriminating specialty coffees produced at different altitudes and under different processing methods.

Following this argument, the authors analyzed a sensory experiment with four specialty coffees produced in the Serra da Mantiqueira region of Minas Gerais, Brazil. Statistical modeling based on the distribution of extreme values was used, considering the highest sensory scores as random phenomena, since there may be variations in the consumers' judgment due to external factors, such as fatigue, sensory ability, among others.

To better understand these scores, the theory of extreme values plays a relevant role in the study of atypical, rare, low-frequency events or events that are occasionally discriminated as outliers. This theory consists of two methodologies: one in which extreme events occur in blocks, which are modeled by the generalized extreme value (GVE) distribution; and another one in which extreme events are defined as those that surpass a threshold, also known as peaks over threshold (POT) (THOMAS *et al.*, 2016). However, given the complexity of integral resolution, analytical results are obtained using a statistical model.

Plausible approximations have been obtained by using computationally intensive methods, such as importance sampling. In this method, a new distribution function is introduced and the values are corrected by weights, thus preventing changes to the expected results.

Given the above, this study used the importance sampling technique along with Monte Carlo simulation as a tool to evaluate probabilistic models based on the distribution of extreme values to model the sensory quality of four specialty coffees produced in the Serra da Mantiqueira region.

MATERIAL AND METHODS

Data on the scores assigned to each coffee type in the sensory experiment were obtained from tests performed at the Federal University of Lavras (Universidade Federal de Lavras - UFLA). In compliance with the decision awarded by the Ethics Committee, as registered in the Certificate of Presentation for Ethical Consideration (CAAE): 14959413.1.0000.5148, the samples of coffee arabica were prepared by removing defective grains and roasted, respecting the minimum period of 24 hours before tasting, according to the protocols of the Specialty Coffee Association of America.

The roasting point was visually determined using the SCAA/Agtron Roast Color Classification System with

Table 1 - Description of the specialty coffees evaluated in the sensory analysis with untrained consumers

Type	Genotype	Altitude	Processing
A	Yellow Bourbon	above 1,200 m	Natural
B	Acaiá	below 1,100 m	Peeled cherry
C	Acaiá	below 1,100 m	Natural
D	Yellow Bourbon	above 1,200 m	Peeled cherry

Fonte: Liska *et al.* (2015)

standard color wheels, following Ferreira *et al.* (2016). The beverage was prepared at a concentration of 7% w/v using filtered water ready for consumption, free of contaminants and without added sugar. With these specifications, four specialty coffee types, whose samples were coded A, B, C and D, were prepared and are described in Table 1.

The four coffee types were evaluated as to their sensory characteristics: taste, acidity, body and note. In different sessions, the volunteering consumers were grouped into two classes: (a) individuals who are used to consuming coffee, but lack basic knowledge on specialty coffees and (b) individuals who are used to consuming coffee and have been provided with basic information on specialty coffees.

Using the POT method, the notes above the pre-established threshold were considered. Considering that a generalized distribution can also be defined using the surpassing values method, we assume that when X is a random variable corresponding to the extreme value probability function, and given a normalized threshold u , the variable $X - u$, which represents the extreme values, follows a generalized Pareto distribution (GPD).

$$P_{\gamma, \beta(u)}(\gamma) = \begin{cases} 1 - \left(1 + \gamma \frac{\gamma}{\beta(u)}\right)^{-\frac{1}{\gamma}}, & \text{for } \dots \gamma \neq 0, \\ 1 - e^{-\frac{\gamma}{\beta(u)}}, & \text{for } \dots \gamma = 0 \end{cases} \quad (1)$$

in which, $\beta(u) > 0, 0 \leq \gamma$ when $\gamma \leq \gamma$ e $\gamma \leq -\frac{\beta(u)}{\gamma}$ when $\gamma < 0$.

Through this distribution, specifying the value assumed by the γ parameter results in the distributions described in Table 2.

Thus, to analyze the chosen threshold with the aid of the POT package (RIBATET, 2007) of software R (R DEVELOPMENT CORE TEAM, 2014), mean residual life plots and plots of the parameters estimated as a function of the threshold were used, and using the MASS and EVD packages of R, the distributions for the sensory notes were fitted using the maximum likelihood method (COLES, 2001).

The development of the importance sampling method starts with the estimation of the expected value, defined by $E[X]$ and obtained by solving the integral.

$$I = \int_0^x \chi f(x) dx. \quad (2)$$

Using the relationship between *Uniform* distribution (0,1) and GPD (μ, σ, ξ), considering $\xi \neq 0$ and the random variable $U \sim Uniforme(0,1)$, then $X = \mu + \sigma(U^{-\xi} - 1)/\xi \sim GPD(\mu, \sigma, \xi \neq 0)$ Assuming $\xi = 0$, then $x = \mu - \sigma \ln(U) \sim GPD(\mu, \sigma, \xi = 0)$. Therefore, with these new specifications, equation (2) has been rewritten according to expression (3).

$$I = \int_0^1 \mu f_u(u) du. \quad (3)$$

By the relationship between, the inverse transformation is given by:

$$X = u + \frac{\sigma(U^{-\xi} - 1)}{\xi} \Rightarrow X = \left[\xi \left(\frac{x - u}{\sigma} \right) + 1 \right]^{-\frac{1}{\xi}}, \text{ if } \xi \neq 0$$

$$X = u - \sigma \ln(U) \Rightarrow U = e^{-\frac{x - u}{\sigma}}, \text{ if } \xi = 0$$

Therefore, integral in (3) can be written as:

$$I = \int_0^1 \mu f_u du = \int_0^1 \exp\left\{-\frac{x - u}{\sigma}\right\} f_u \left(\exp\left\{-\frac{x - u}{\sigma}\right\} \right) dx. \quad (4)$$

Nevertheless, the integral in (4) is in the scale of *Uniform* distribution, hence, to return to GDP distribution, which is given by:

$$X^* = u + \frac{\sigma^* (U^{*-\xi} - 1)}{\xi}, \text{ if } \xi \neq 0$$

$$X^* = u - \sigma \ln(U^*) \text{ if } \xi = 0.$$

To performing importance sampling in this research study, we considered candidate distribution $h(x)$ given by $N(u; \sigma=1)$, fixed, with parameters u and σ of $GDP(u; \sigma; \xi)$. The $h(x)$ must be chosen so that probability will not be equal to zero, otherwise it will result in infinite weights. Thus, Normal distribution focused on u and $\sigma = 1$ as $h(x)$ becomes a convenient choice because it finds support in $(-\infty, +\infty)$.

Therefore, I in (4) was solved following (5).

$$I = \int_0^1 \exp\left\{-\frac{x - u}{\sigma}\right\} \frac{f_u \left(\exp\left\{-\frac{x - u}{\sigma}\right\} \right)}{\phi_{\mu, \sigma}(x)} \phi_{u, \sigma}(x) dx = \int_0^1 g^*(x) \phi_{u, \sigma}(x) dx \quad (5)$$

Whose solution is obtained by resampling $g^*(x^*)$.

$$\hat{I} = \frac{1}{n} \sum_{i=1}^n g^*(x_i^*) \quad (6)$$

Table 2 - Distributions corresponding to the values assumed by the γ parameter

Parameter	Distribution
$\gamma=0$	Exponential type I belonging to the Gumbel class
$\gamma>0$	Pareto type II belonging to the Fréchet class
$\gamma>0,$	Pareto type III belonging to the Weibull class

Source: Coles (2001)

in which $n = 1000$ performing Monte Carlo simulations and x_i^* the realization of the random variable X_i^* .

The fit adequacy of each distribution was validated using the Kolmogorov-Smirnov (KS) goodness-of-fit test, which verifies the fit of a probability distribution to the original data. Further information on this test can be found in Thomas *et al.* (2016).

The Ljung-Box test was used to verify the assumption of independence of the observations. More details can be found in Ljung and Box (1978). In both tests, a level of significance of 1% ($p < 0.01$) was adopted.

RESULT AND DISCUSSION

Importance sampling was used to provide an approximation for $E[X]$, with X , the random variable, associated with the maximum note provided by a taster. As previously noted, X follows a GDP.

The results were organized into two situations: the first refers to importance sampling obtained from the original sample; and the second considers the fact that X follows a GDP. In the second case, GDP parameter estimations fitted to the original data and sampled values of this distribution were employed to the importance sampling.

Table 3 presents the results of fitting the GDP to the different specialty coffee types for the following parameters: μ corresponding to mean, σ to scale and ξ to shape. Coffee D had the highest expected value for the maximum notes, and on average, the maximum note attributed to this coffee was 9.7 points. In contrast, coffee C had the lowest expected value for the maximum note.

Another expressive result is verified regarding the Ljung-Box and Kolmogorov Smirnov tests, which allow to interpret that the maximum sensory notes given by each volunteer can be considered independent and satisfactorily fit by the GDP at the 1% significance level.

Briefly, importance sampling was employed to estimate $E[X]$, and the results obtained are reported in Table 4. For that purpose, the relationship between the uniform distribution and the GDP was used to obtain $f(x)$, and the choice of the candidate distribution for $h(x)$ was made such that $h(x)$ did not provide probabilities equal to zero. According to this justification, the normal distribution centered on μ and with variance 1 was used as the candidate distribution.

The means obtained by importance sampling for the first situation are quite close to the theoretical values, except for the mean obtained for coffee B, whose estimate was 9.0 points, whereas the one obtained by importance sampling was 7.5 points.

Considering the sampling performed in the second situation, very reasonable approximations were also observed for the theoretical mean. For example, for coffee D, the theoretical mean was 9.7 points, and the approximation via importance sampling was 9.8 points. For coffee B, the approximation for $E[X]$ did not show good accuracy because the average Monte Carlo for $E[X]$ was discrepant in relation to the theoretical value. Important additional information is the precision of the importance sampling, given by the Monte Carlo standard deviation of the simulation, which indicates the variability of the samples generated by $f(x)$. Moreover, coffee B had the lowest precision among the specialty coffees, indicating that for this coffee, the notes presented greater variability in relation to the mean of the GDP, which is confirmed by the estimated reported in Table 4.

Considering the results in Table 3, importance sampling was successfully employed for coffees A, C and D, and considering the samples from the GDP, it was possible to establish a precision measure for the expected value of the maximum notes for these coffees.

Figures 1, 2, 3 and 4 corroborate the results of the importance sampling, as the behavior of X is symmetric, as can be observed from the histograms.

Table 3 - Results of the fit of the generalized Pareto distribution, Ljung-Box test (Q), likelihood ratio test, Kolmogorov-Smirnov test (KS) and expected value of the random variable $E[X]$

Coffee		coffee A	coffee B	coffee C	coffee D
Parameters	μ	8.1000	8.2000	8.0000	9.4000
	σ	0.5314	1.5019	0.1075	0.6198
	ξ	0.0000	-0.8566	0.0000	-1.0327
Q		0.4399	0.3325	0.1810	0.1150
Ratio		0.0294	0.0002	0.0269	<0.0001
KS		0.2974	0.1618	0.4262	0.2069
$E[X]$		8.6314	9.0089	8.5481	9.7049

Table 4 - Results of the importance sampling (IS) considering the original sample and the sample from a GPD, as well as the Monte Carlo precision

Coffee		coffee A	coffee B	coffee C	coffee D
Theoretical	$E[X]$	8.6314	9.0089	8.5481	9.7049
IS Considering original sample	$E[X]$	8.2534	7.5166	8.1476	9.9436
IS considering GP distribution sample	$E[X]$	8.2077	7.5185	8.1021	9.851
	SD	0.0593	0.1056	0.0544	0.062
	$h(x)$	$N(\hat{\mu}, 1)$	$N(\hat{\mu}, 1)$	$N(\hat{\mu}, 1)$	$N(\hat{\mu}, 1)$

Figure 1 - Histogram generated from 1000 Monte Carlo simulations with sensory notes for coffee A

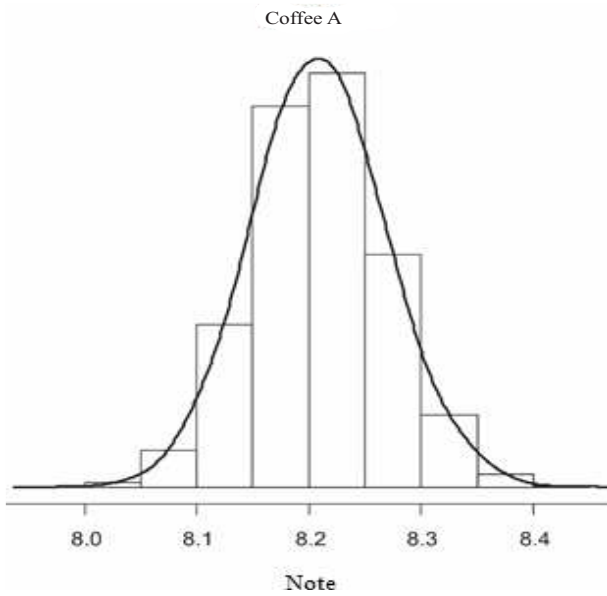


Figure 2 - Histogram generated from 1000 Monte Carlo simulations with sensory notes for coffee B

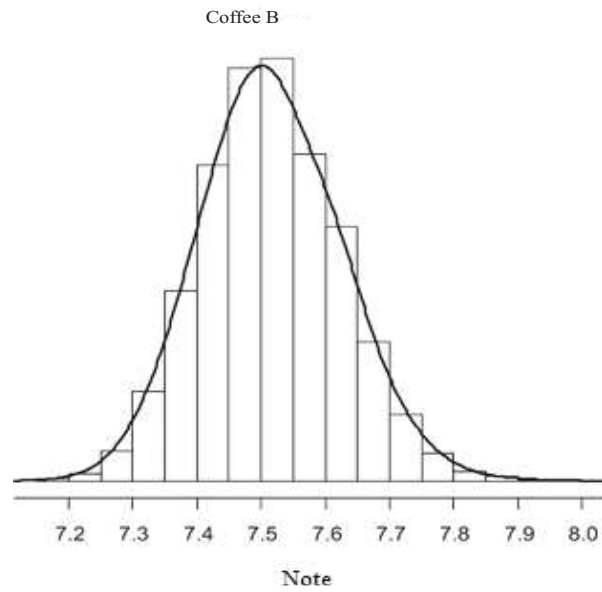


Figure 3 - Histogram generated from 1000 Monte Carlo simulations with sensory notes for coffee C

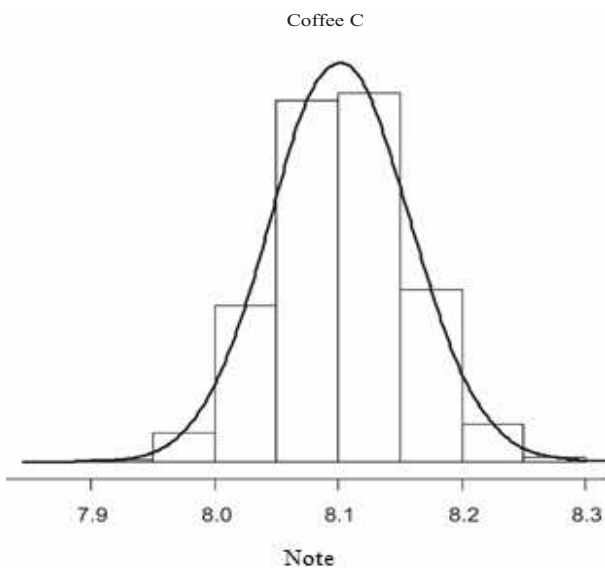
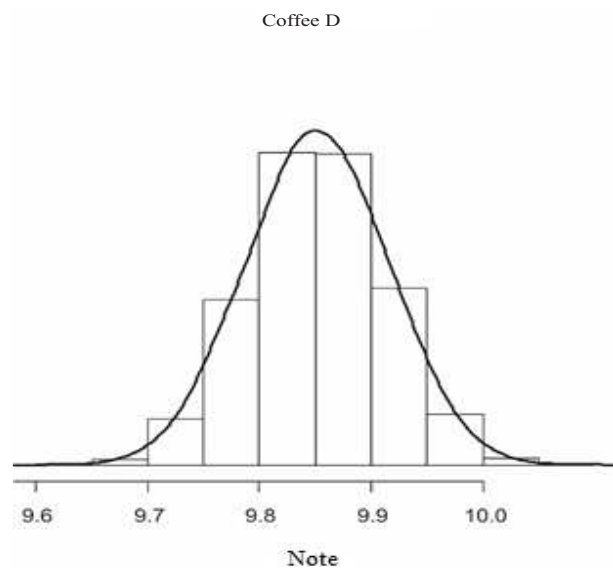


Figure 4 - Histogram generated from 1000 Monte Carlo simulations with sensory notes for coffee D



The importance sampling method allows some flexibility in the choice of the candidate distribution for $h(x)$, and suggests that the one that behaves similarly to the desired distribution, in this case the GPD. For example, the Gumbel and Weibull distributions were used, but none achieved results better than those reported in Table 4. In this sense, further studies can be performed to improve the results reported in Table 4, especially those for coffee B.

CONCLUSIONS

1. The GP distribution can be applied to the sensory analysis of specialty coffees made by a heterogeneous sensory panel of consumers;
2. The importance sampling method was successfully used for the specialty coffees of the *Coffea arabica* species genotypes Yellow Bourbon and Acaiá. Coffee type D had the highest Monte Carlo mean sensory note and high Monte Carlo precision. Coffee type B had the highest variability among the analyzed coffees.

REFERÊNCIAS

- BORÉM, F.M. *et al.* The relationship between organic acids, sucrose and the quality of specialty coffees. **African Journal of Agricultural Research**, v. 11, p. 709-717, 2016.
- CHAGAS, E. N. *et al.* Selection of robust estimators used in analysis of sensory characteristics and identification of environments conducive to specialty coffee production. **Advanced Crop Science**, v. 3, p. 515-524, 2013.
- COLES, S. **An introduction to statistical modeling of extreme values**. 3. ed. London: Springer, 2004. 221 p.
- FERREIRA, H. A. *et al.* Selecting a probabilistic model applied to the sensory analysis of specialty coffees performed with consumer. **IEEE Latin America Transactions**, v. 14, n. 3, 2016.
- LISKA, G. R. *et al.* Evaluation of sensory panels of consumers of specialty coffee beverages using the boosting method in discriminant analysis. **Semina: Ciências Agrárias**, v. 36, n. 6, p. 3671-3680, 2015.
- LJUNG, G. M.; BOX, G. E. P. *On a measure of lack of fit in time series models*. **Biometrika**, v. 65, p. 297-303, 1978.
- MALTA, M. R.; CHAGAS, S. J. R. “Avaliação de compostos não voláteis em diferentes cultivares de cafeeiro produzidas na região sul de Minas Gerais”. **Acta Scientiarum Agronomy**, v. 31, n. 1, p. 57-61, 2009.
- OSSANI, P. C. *et al.* Qualidade de cafés especiais: uma avaliação sensorial feita com consumidores utilizando a técnica MFACT. **Revista Ciência Agrônômica**, v. 48, n. 1, p. 92-100, 2017.
- R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2014.
- RAMOS, M. F. *et al.* Discrimination of the sensory quality of the *Coffea arabica* L. (cv. Yellow Bourbon) produced in different altitudes using decision trees obtained by the CHAID method. **Journal of the Science of Food and Agriculture**, v. 96, n. 10, p. 3543-3551, 2016.
- RIBATET, M. *A user's guide to the POT Package (version 1.4)*. Québec: University of Québec, 2007.
- RIBEIRO, D. E. *et al.* Interaction of genotype, environment and processing in the chemical composition expression and sensorial quality of Arabica coffee. **African Journal of Agricultural Research**, v. 11, p. 2412-2422, 2016.
- THOMAS, M. *et al.* Applications of Extreme Value Theory in Public Health. **PLOS ONE**, v. 11, n. 7, 2016.