We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Intensive Computational Method Applied for Assessing Specialty Coffees by Trained and Untrained Consumers

Gilberto Rodrigues Liska, Luiz Alberto Beijo,
Marcelo Ângelo Cirillo, Flávio Meira Borém
and Fortunato Silva de Menezes

Abstract

The sensory analysis of coffees assumes that a sensory panel is formed by tasters trained according to the recommendations of the American Specialty Coffee Association. However, the choice that routinely determines the preference of a coffee is made through experimentation with consumers, in which, for the most part, they have no specific ability in relation to sensory characteristics. Considering that untrained consumers or those with basic knowledge regarding the quality of specialty coffees have little ability to discriminate between different sensory attributes, it is reasonable to admit the highest score given by a taster. Given this fact, probabilistic studies considering appropriate probability distributions are necessary. To access the uncertainty inherent in the notes given by the tasters, resampling methods such as Monte Carlo's can be considered and when there is no knowledge about the distribution of a given statistic, p-Bootstrap confidence intervals become a viable alternative. This text will bring considerations about the use of the nonparametric resampling method by Bootstrap with application in sensory analysis, using probability distributions related to the maximum scores of tasters and accessing the most frequent region (mode) through computational resampling methods.

Keywords: probability, Monte Carlo, bootstrap, GEV distribution, Mantiqueira Serra, height, consumers

1. Introduction

Basically the methodology involved in the analysis of sensory data is summarized in a set of experimental and statistical techniques applied with the purpose of verifying the quality or the degree of acceptance of a given product, without, however, disregarding the characteristics of the individuals, with respect to your sensory skills. In this context, two distinct groups of consumers can be inserted, that is, consumers who have some enhanced sensory ability (s), resulting from product training or knowledge and totally lay consumers.

Faced with this situation, it becomes plausible to admit that a sensory analysis, applied to a group of trained consumers, being able to discriminate small differences between the samples, the results provided by the evaluations will show little variation [1]. Therefore, a sensory experiment carried out with this group shows a greater agreement with the procedures standardized by [2], since the objective assessments would be more homogeneous for the perception of uniformity, sweetness, defects, among others, mentioned by [3, 4].

In an opposite situation, considering a group of untrained consumers, it is more likely that the evaluations will present heterogeneous results, in such a way that the statistical treatment to be given in the analysis of these results may include the atypical observations, classified as outliers arising from the evaluation. Individual to each consumer [5, 6].

It is worth mentioning that the heterogeneity between the observations may be the result of uncontrollable factors, such as, for example, genetics, fatigue, unwillingness to carry out all tests and differences between the abilities of consumers, as well as external causes such as, for example, the geographical origin of a particular product whose qualities or characteristics are due exclusively or essentially to the geographical environment, including natural and chemical factors, which, among others, mention variations in chemical composition due to the genetic variability between cultivars that influence the sensory quality of coffees [7–12].

Given countless causes that are supposed to be the sources that cause outliers in a sensory analysis and reporting the analysis of the quality of coffees, special coffees can be highlighted. Following the definition given by [2], in summary, a coffee is said to be special, as it presents superior quality to its competitors in relation to its origin, absence of defects, processing and/or sensory expressions such as aroma, flavor.

The results of the sensory evaluation are established on a scale ranging from 0 to 10 in which these values represent the increasing levels of coffee quality. According to the analysis protocol [2] the results of the sensory evaluation vary according to a scale where the grades 6, 7, 8, 9 correspond respectively to: good, very good, excellent and exceptional. When the grades are less than 6, the coffees are declared to be of a quality below the Specialty Grade.

Respecting these characteristics, Coffee arabica cultivars are potential coffees worthy of being classified as special [13, 14]. However, studies related to the interference of the environment and geographic origin can influence the quality of the drink. [14], in a study interacting quality with environmental factors, concluded that the coffees with the highest scores in a contest held in the state of Minas Gerais, were produced in colder regions with milder temperatures and annual precipitation index around 1600 mm [15]. In this context, in humid regions it is recommend that processing be performed prioritizing peeled and desmucilated coffees. Thus, the quality of the coffee would be inferred without the interference of defects.

In the case of statistical methodology, it is highlighted that the usual methods of analysis, in general, are sensitive to outlier observations, these being plausible to have arisen in a sensory analysis carried out by untrained consumers [5, 15].

Due to this fact and assuming that the assignment of maximum sensory scores can be understood as random phenomena, in the sense that there are variations in the judgment of different consumers, this work aims to propose the use of some distributions belonging to the generalized extreme value distribution class in sensory analysis. For this purpose, this work analyzes a sensory experiment to evaluate four special coffees produced in the Serra da Mantiqueira Region of Minas Gerais, differentiated in preparation and geographical identification classified by different altitudes.

Bootstrap, developed by Efron in the 70s, can be used in many situations. It is based on a simple, yet powerful idea that the sample represents the population, so

analogous characteristics of the sample should give us information about the characteristics of the population. Bootstrap helps to learn about these sample characteristics by taking resamples (samples with replacement of the original sample) and we use this information to infer about the population [16].

In this sense, to detect a difference in the judgment of special coffees by trained and untrained tasters, a test built via non-parametric Bootstrap will be proposed for the mode of distribution of extreme values that best fits the data set.

2. Modeling maximum sensory scores and numerical procedure

In accordance with the opinion of the Ethics and Research Council, registered with the CAAE: 14959413.1.0000.5148, the preparation of the Samples of 100% Arabica coffee was done by removing all defective beans and toast, respecting the maximum period of 24 hours for tasting.

The roasting point was determined visually, using the color classification system by means of standardized discs (SCAA/Agtron Roast Color Classification System). Regarding the preparation of the drink, the concentration of 7% w/v was maintained using filtered water ready for consumption, free of any contaminants and without added sugar. With these specifications, four types of specialty coffees, coded in the samples by A, B, C and D given the description in **Table 1**.

For each type of coffee, the following sensory characteristics were assessed in the acceptance test: aroma, body, hardness, and final score, in four sessions, with the participation of a volunteer group of consumers with basic knowledge in regard to sensory analysis of coffees and another group without basic knowledge. **Table 2** provides a list of the tasters, as well as the sensory characteristics assessed by each taster, in which a_{ij} represents the score given by taster i ($i = 1, 2, \ldots, n_1, n_1 + 1,$ $n_1 + 2, \ldots, n_2$), such that $n_1 + n_2 = n$, for the sensory characteristic × coffee j ($j = 1, 2, \ldots, 16$) combination.

In the test, four different types of coffee were evaluated in terms of their sensory characteristics, flavor, acidity, body and note. In different sessions, voluntary consumers were grouped into two classes: (a) people with the habit of consuming coffee, but who do not have basic knowledge about specialty coffees and (b) people with the habit of consuming coffee and trained with information basic information about specialty coffees.

The fit of the probability distributions was carried out, considering the random variable X representing the maximum consumers' sensory scores for the each type of coffee (**Table 1**), totaling in a sample of 696 observations.

Bearing in mind that the highest score provided by a tester will be considered, this being considered as a block, the distribution of the maximums, according to the Fisher-Tippet theorem, is the generalized extreme values distribution (GEV). Its probability density function is defined by:

Type	Genotype	Altitude	Processing
A	Bourbon	Above 1200 m	Natural
B	Acaia	Below 1100 m	Pulped natural
C	Acaia	Below 1100 m	Natural
D	Bourbon	Above 1200 m	Pulped natural

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

Condition	Taster	Sensory characteristic 1				...	Sensory characteristic 4			
		A	B	C	D	...	A	B	C	D
Trained	1	a_{11}	a_{12}	a_{13}	a_{14}	...	a_{113}	a_{114}	a_{115}	a_{116}
	2	a_{21}	a_{22}	a_{23}	a_{24}	...	a_{213}	a_{214}	a_{215}	a_{216}
	.	.	.	.	.	...	.	.	.	.
	.	.	.	.	.		.	.	.	.
	.	.	.	.	.		.	.	.	.
	n_1	a_{n11}	a_{n12}	a_{n13}	a_{n14}	...	a_{n113}	a_{n114}	a_{n115}	a_{n116}
Untrained	1	$a_{(n1+1)1}$	$a_{(n1+1)2}$	$a_{(n1+1)3}$	$a_{(n1+1)4}$	...	$a_{(n1+1)13}$	$a_{(n1+1)14}$	$a_{(n1+1)15}$	$a_{(n1+1)16}$
	.	.	.	.	.	...	.	.	.	.
	.	.	.	.	.		.	.	.	.
	.	.	.	.	.		.	.	.	.
	n_2	a_{n21}	a_{n22}	a_{n23}	a_{n24}	...	a_{n213}	a_{n214}	a_{n215}	a_{n216}

Table 2.
Tabulated representation of the sensory characteristics of the specialty coffees assessed.

$$f(x; \mu, \sigma, \xi,) = \frac{1}{\sigma}\left\{\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\left(-\frac{1}{\xi}\right)-1} \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)^{-\left(\frac{1}{\xi}\right)}\right]\right\}\right\}, \quad (1)$$

where $-\infty < x < \mu - \sigma/\xi$ when $\xi < 0$ resulting in the Weibull (μ, σ, ξ). When $\mu - \sigma/\xi < x < \infty$ for $\xi > 0$ results in Fréchet (μ, σ, ξ). When $\lim_{\xi \to 0} f(x; \mu, \sigma, \xi)$ leads to the Gumbel distribution. The parameters μ, σ and ξ are the location, scale and shape parameters.

The probability that a maximum score will be greater than realization of a score, represented by x is defined as

$$P[X > x] = 1 - P[X \le x] = 1 - F(x; \hat{\theta}), \quad (2)$$

where $\hat{\theta}$ corresponds to the vector of maximum likelihood estimates [17, 18]. This method requires that the maximum scores are independent and identically distributed [19], which was assessed by the Ljung Box test, explained follow. $F(x; \hat{\theta})$ is the cumulative distribution function (cdf) of GEV probability density function. Its cdf is given by

$$F(x; \mu, \sigma, \xi) = \exp\left[-\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{\frac{-1}{\xi}}\right]. \quad (3)$$

The mode (*Mo*) of the pdf in Eq. (1) is given by

$$Mo(X) = \mu + \sigma\frac{(1+\xi)^{-\xi} - 1}{\xi}, \quad (4)$$

and in the $\lim_{\xi \to 0} f(x; \mu, \sigma, \xi)$ case, the mode is simplified to $Mo(X) = \mu$.

The goodness of fit for each distribution was validated using the Kolmogorov Smirnov (KS) adherence test in conjunction with the Q-Q plots [20, 21]. The Q-Q plot consists of the points

$$\left\{\left(F^{-1}(p_i), x_i\right), i = 1, \ldots, n\right\}, \quad (5)$$

where $F^{-1}(p_i)$ is the inverse function of the cumulative distribution function of a given probability distribution, p_i are the percentiles and x_i are the data used to fit the model, ordered in ascending and n the sample size.

According to [22], the Kolmogorov–Smirnov (KS) test is used to assess the fit of a probability distribution to the original data. It is based on the analysis of the proximity or adjustment between the sample distribution function $\hat{F}(x_i)$ and the population distribution function under the null hypothesis, $F_0(x_i)$. The test statistic (D) is given by,

$$D = \max |F_0(x_i) - \hat{F}(x_i)| \qquad (6)$$

In the KS test, the hypothesis of interest are given by H_0: The distribution function from which the sample is derived follows the distribution function that is assumed to be known; that is, $F(x) = F_0(x_i)$ and H_1: $F(x) \neq F_0(x_i)$ [23]. Then, the value (Eq. (6)) must be compared with the critical value (using tables), for the significance level of the test. According on the result, the null hypothesis is rejected or not. The null hypothesis is also rejected if the p-value is lower than the significance level adopted.

Regarding the verification of the assumption of independence of the observations, such that is required by the maximum likelihood method for estimating parameters, the Ljung-Box (LB) test was used. According to [24], it is a statistical test used to find out if there are non-zero autocorrelation groups. To do this, it tests total randomness based on the number of deviations. The test hypotheses are H_0: all autocorrelation coefficients are equal to zero and H_1: not all autocorrelation coefficients are equal to zero. The test statistic is

$$Q = n(n+2) \sum_{k=1}^{s} \frac{r_j^2}{(n-j)}, \qquad (7)$$

where n is the number of observations, s is the number of coefficients in testing autocorrelation, r_j is the autocorrelation coefficient (for the deviation) and Q the test statistic. If the sample values of Eq. (7) exceed the critical value of a Chi-Squared distribution with s degrees of freedom, then at least one deviation r is statistically different from zero at the specified significance level, that is, H_0 is rejected. H_0 is also rejected if p-value is lower than the adopted significance level. It should be noted that if H_0 is rejected, it can be said that the data are independent. In both tests, the significance level of 1% was adopted [25].

In order to make an inference about the most frequent score among the tasters, it is necessary to know the sample distribution of the quantity in Eq. (4). For that, an alternative would be to use resampling methods, which one of them will be presented below.

The Bootstrap resampling process consists of resampling B samples $P^{*(1)}, P^{*(2)}, \ldots, P^{*(B)}$, with replacement, independent and identically distributed of the n highest marks awarded by trained and untrained tasters. Estimates of the parameter of interest can be obtained, denoted by $\hat{\theta}^*_{(i)}$, for each sample, which is $\theta = Mo(X)$. With that we will obtain the vector $\hat{\boldsymbol{\theta}}^* = \left(\hat{\theta}^*_{(1)}, \hat{\theta}^*_{(2)}, \ldots, \hat{\theta}^*_{(B)} \right)$ and from the vector $\hat{\boldsymbol{\theta}}^*$ it is possible to obtain the Bootstrap distribution of the $\hat{\theta}$ estimator.

Once the empirical distribution of the $\hat{\theta}$ estimator is obtained, confidence intervals for θ can be estimated. The Bootstrap confidence interval based on the Bootstrap distribution percentiles of θ, described in [16, 26], is known as the p-Bootstrap confidence interval. In a more formal way, the confidence interval can be constructed by following the following steps:

(**Step 1**) Draw, with replacement, of P, one Bootstrap sample P^*;

(**Step 2**) From Bootstrap sample P^*, obtain $\hat{\theta} = Mo(X)$;

(**Step 3**) Repeat the steps 1 and 2 B times;

(**Step 4**) From the vector $\hat{\boldsymbol{\theta}}^* = \left(\hat{\theta}^*_{(1)} \leq \hat{\theta}^*_{(2)} \leq \ldots \leq \hat{\theta}^*_{(B)} \right)$, for α significance level $(0 < \alpha < 1)$, the p-Bootstrap confidence interval with $100 \times (1 - \alpha)\%$ level of confidence is given by $IC_{(1-\alpha)}(\theta) : \left[\hat{\theta}^*_{(k_1)}; \hat{\theta}^*_{(k_2)} \right]$, where $k_1 = (B+1)(\alpha/2)$ and $k_2 = (B+1)(1 - \alpha/2)$ are the highest integers that are not greater than$(B+1)(\alpha/2)$ and $(B+1)(1 - \alpha/2)$, respectively; and $\hat{\theta}^*_{(k_1)}$ is the $100(\alpha/2)\%$-percentile of the Bootstrap empirical distribution; and $\hat{\theta}^*_{(k_2)}$ is the $100(1 - \alpha/2)\%$-percentile of the Bootstrap empirical distribution [16, 26].

Finishing the proposed methodology, the computational resources available in the R software [27, 28] were used through the *boot* and *evd* [29] packages to fitting the probability distributions for sensory scores, hypothesis tests and construction of Bootstrap confidence intervals.

3. Experimental results

The following results correspond to the parameter estimates for the probability distributions fitted for the two classes of tasters, as well as the p-values referring to the validation of the probabilistic model fitted for the sensory scores.

With these specifications, given a level of significance of 1%, it is noted the confirmation of the fit in the sensory scores for each coffee, therefore, there is statistical evidence to assume that GEV distribution is adequate to model the maximum sensory grades of the evaluated coffees (**Table 3**). It should be noted that the fact that we have p-values greater than 1% for the KS test indicates that there is statistical evidence for the acceptance of the test's null hypothesis, as can be seen in Section 2. The test used, however, according to [30], should only be used for completely specified distributions, that is, when there are no unknown parameters that need to be estimated from the sample. Otherwise, the test is very conservative. One solution would be to obtain, via simulation, the theoretical quantiles of the Kolmogorov Smirnov test to compare them with the quantiles obtained from the sample. A similar procedure for the Gumbel distribution was carried out by [31].

Coffee	Group	Parameter estimates			KS (*p*-value)	LB (*p*-value)
		$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$		
A	Untrained	5.9471	2.4105	-0.6569	0.9077	0.0803
	Trained	6.9345	1.6259	-0.6156	0.8908	0.2359
B	Untrained	5.9326	2.4624	-0.5721	0.9466	0.3306
	Trained	6.6031	1.9455	-0.5736	0.8255	0.0110
C	Untrained	6.4290	2.2108	-0.6348	0.9485	0.6084
	Trained	7.0595	1.3382	-0.5485	0.9998	0.9823
D	Untrained	7.8676	2.0437	-0.9582	0.9962	0.9625
	Trained	7.8113	1.8183	-0.8221	0.6543	0.6924

Table 3.
Parameters estimates and results of the Kolmogorov–Smirnov and Ljung-box tests for the maximum scores given by consumers in the sensory evaluation of the special coffees named in A, B, C and D.
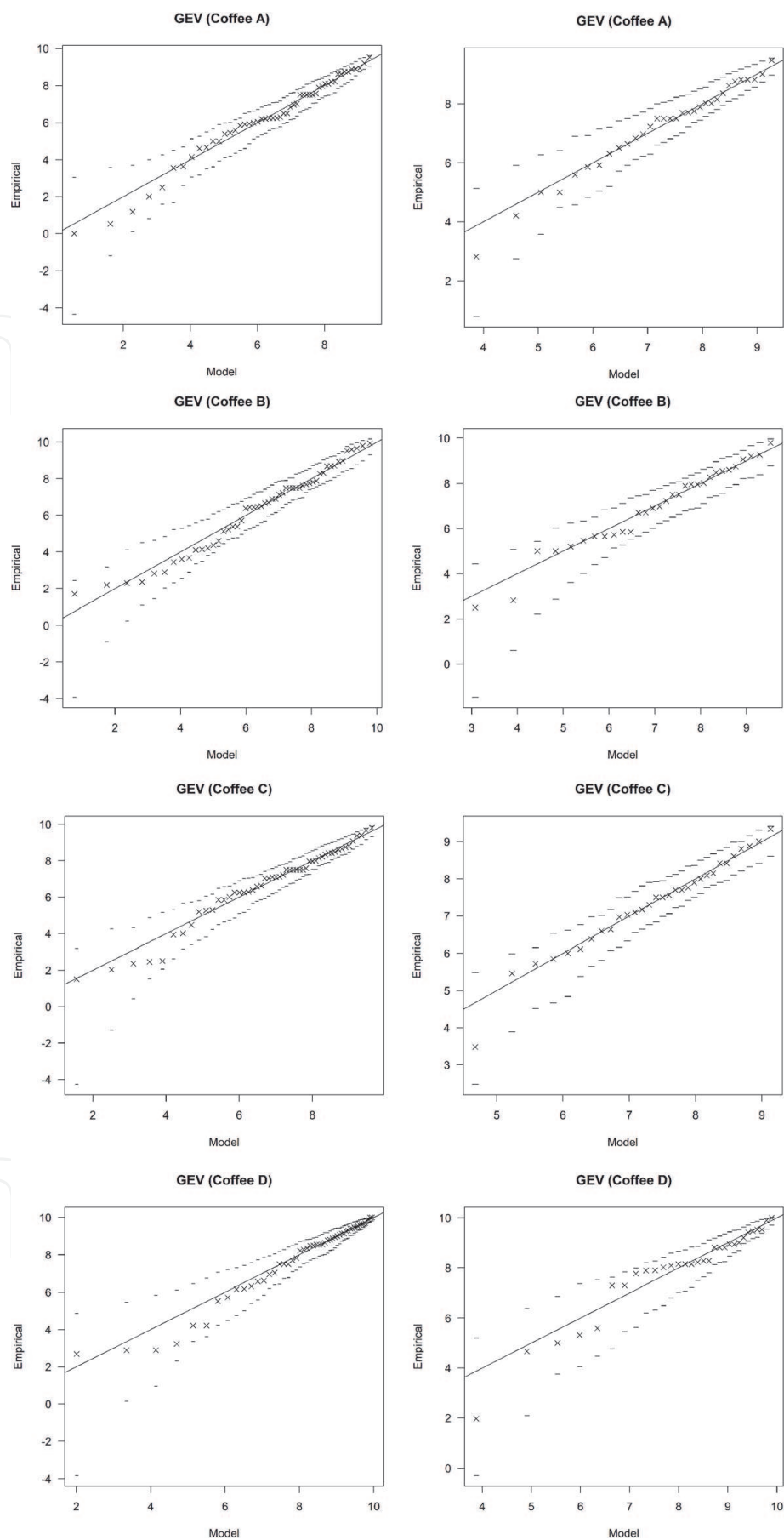
Figure 1.
Q-Q plot referring to the fitted GEV distributions for the maximum sensory scores obtained in the evaluation of each special coffee for the group of untrained (left) and trained (right) tasters.

Alternatively, inspection of fit quality can be assessed via Q-Q plots graphs. They are shown in **Figure 1**.

In this sense, the validation of the GEV distribution is corroborated in the Q-Q plots shown in **Figure 1**, because for all the specialty coffees evaluated, the theoretical quantiles showed a linear behavior and close to the straight identity with the observed quantiles and the points being, in their mostly, contained in the 95% confidence interval. It should also be noted that the quantiles have a trend to converge to a region located as an upper tail. All p-values of the Ljung Box test are higher than 1%, thus showing the acceptance of the null hypothesis of the test, as described in Section 2. It can be concluded, therefore, that the maximum scores given by trained and untrained tasters they are independent. We should highlight that we have used these tests to verify the assumptions of the Extreme Value Theory models, but that they could be used for other interests, such as in the trend analysis of hydro-climatic series [32–34]. Failure to observe these assumptions can lead to fitted models parameter estimates, as well return levels estimates, biased and/or under/overestimated. For these situations, Bayesian methods, regression or time series based on the Box-Jenkins methodology could be considered [35, 36].

In function of the confirmatory results related to the GEV distribution goodness of fit, given the estimates of the parameters for this distribution applied in the maximum sensory scores given by consumers in the evaluations carried out for each coffee, we proceeded with the calculations of the probabilities for an individual to supply a grade higher than a given grade. The results are described in **Table 4**.

Before that, the distribution modes were calculated as shown in **Table 4**, in order to verify the similarity between the grades provided by trained and untrained tasters. It is observed that occasionally they can be considered very close. For specialty coffees A and B, trained tasters provided higher grades more frequently than untrained tasters and for specialty coffees C and D the opposite occurred.

Although the similarity between the modes of the grades attributed by the tasters is evident, this similarity is not associated with any level of confidence, since the similarity is only punctual. To circumvent this situation, confidence intervals were constructed using the non-parametric Bootstrap method, as shown in Section 2. Thus, it can be stated with 95% confidence that the grades most frequently attributed to coffee A by trained tasters and not trained do not differ statistically, since the point estimate for the fashion of the notes is contained in the respective confidence intervals and they are overlapping.

Coffee	Group	Mode	$q_{0.025}$	$q_{0.975}$	$P[X > q_{0.025}]$ (%)	$P[X > q_{0.975}]$ (%)	Difference (%)
A	Untrained	7.8	6.9	8.9	47.2	7.7	39.5
	Trained	8.1	7.3	9.0	54.1	8.2	45.8
B	Untrained	7.6	6.2	9.2	59.2	8.2	51.0
	Trained	7.9	6.6	9.2	62.4	7.1	55.4
C	Untrained	8.1	7.2	9.0	47.6	9.5	38.1
	Trained	7.9	7.1	8.9	62.2	8.1	54.1
D	Untrained	9.9	9.1	10.0	32.8	0.2	32.6
	Trained	9.5	8.6	10.0	44.5	0.7	43.9

*The probabilities P[X > q0.025], P[X > q0.975] and Difference are given in percentages.*

Table 4.
Maximum scores modes given by consumers in the sensory panel of specialty coffees named in A, B, C and D and their respective 95% confidence intervals ($q_{0.025}$ and $q_{0.975}$).

More specifically speaking, for coffee A, the initial mode estimate is 7.8 points for untrained tasters, i.e., $\hat{\theta}_0 = Mo(X)$ is 7.8 points, that is contained in the 95% confidence interval for Bootstrap mode ($\hat{\theta}^*_{(2.5\%)} = 6.9$ points and $\hat{\theta}^*_{(97.5\%)} = 8.9$ points). Likewise for trained tasters, the initial estimate for mode is 8.1 points for trained tasters, that is, $\hat{\theta}_0 = Mo(X)$, is 8.1 points, that is contained in the 95% confidence interval for Bootstrap mode ($\hat{\theta}^*_{(2.5\%)} = 7.3$ points and $\hat{\theta}^*_{(97.5\%)} = 9.0$ points), indicating, therefore, that the scores attributed to coffee A by trained and untrained tasters are similar with 95% confidence.

According to the results described in **Table 4**, it is clear that given a sensory panel made up of untrained consumers, there is a probability that all consumers will have a sensory score higher than 6.0, indicating that whatever the taster is, among the types of specialty coffees studied, no coffee will be classified with quality below the Specialty Grade, since all the coffees analyzed showed a high probability that the most frequent grade is higher than 6. On a 9-point verbal hedonic scale, it can be concluded that, in general, consumers have a trend to be indifferent to the agradability of specialty coffees.

Figure 2 presents the histogram and the Q-Q plot for the mode of the fitted distribution for the grades given by the untrained tasters for coffee A in **Table 1**. The histogram suggests that the empirical distribution of $\theta = Mo(X)$ is a normal and this fact is corroborated by the Q-Q plot, since the one-to-one proportionality is maintained considering the quantiles of the standard normal versus the observed quantiles. Similar results were observed for all other specialty coffees, however these results will not be shown.

When considering an expressive score worthy of international competitions, having a reference higher than 8, the probability of a consumer providing an occurrence of a note being higher than 8 or the coffee being classified as excellent is relatively low for all evaluated coffees (**Table 4**). It is also noted that the probability of a consumer assigning a grade between 9.1 and 10.0 is 32.8%, that is, it can be interpreted that coffee D to be considered exceptional by a consumer is 32.8%. In addition, coffee D is the one with the least amplitude in probability, corresponding to the column "Difference" in **Table 4**, which indicates that it is a type of coffee that provided low variability between the grades attributed by the tasters. On the other hand, coffee B showed greater variability between the grades attributed by
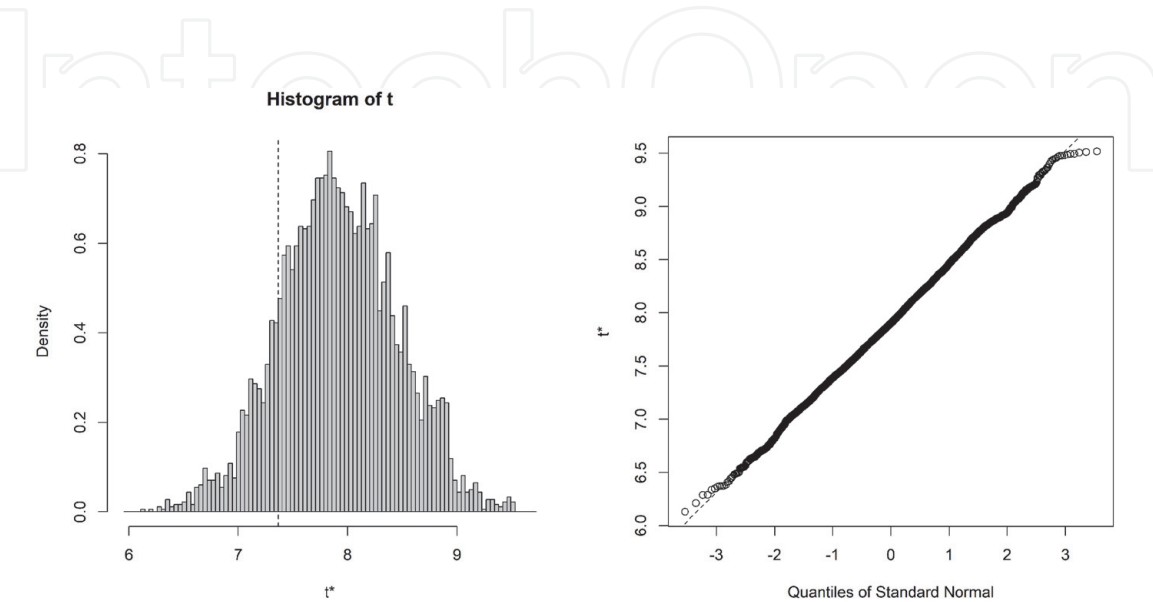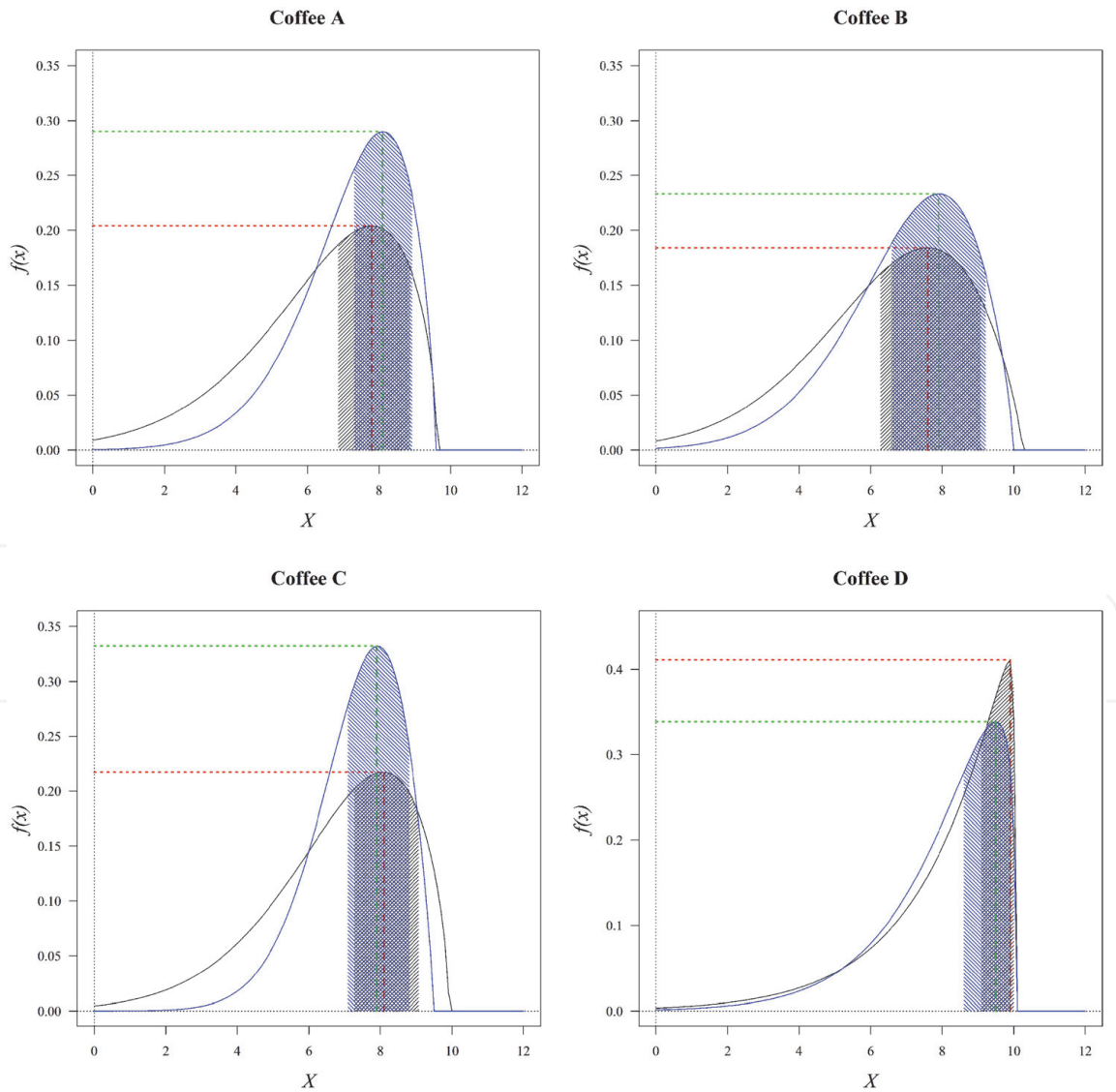

Figure 2.
Histogram for the 5000 values of the bootstrap modes for the scores of untrained tasters and the respective normal Q-Q plot.

the tasters, since the difference $P[X > q_{0.025}]$ - $P[X > q_{0.975}]$ is the largest among the analyzed coffees.

Therefore, in the evaluation of the four specialty coffees, given the low probabilities, it can be said that a sensory experiment carried out with the objective of discriminating the specialty coffees, is done with consumers who present more improved training.

Figure 3 shows graphically the agreement between the scores given by trained (blue hatched) and untrained (black hatched) tasters, according to the results shown in **Table 3**.

The importance of using bootstrap procedures in the analysis of responses that corroborate with these scores is relevant for statistically validating the scores obtained in international competitions, since it assumes that subjective and / or unknown factors, related to the different sensory perceptions of the tasters may suggest violations in the sample distribution, and as a consequence, the estimates of the probabilistic model are distorted. Thus, through successive resampling, an empirical distribution for each parameter is generated in connection with the assumed probabilistic model, and inferences will be made with better precision and accuracy. The amplitude of the confidence interval in **Figure 3** reflects the precision

Figure 3.
Graphical representation of the probability distributions adjusted for the notes attributed by untrained (black) and trained (blue) tasters and the respective 95% confidence intervals for the maximum scores modes attributed to each special coffee.

of the estimates for the maximum notes mode, given the GEV distribution. Other confidence intervals via bootstrap could be considered, such as bootstrap-*t* and BCa [37, 38]. We emphasize that the strategy adopted is innovative in the context of sensory notes and the comparison of confidence intervals can be done as future work.

4. Conclusions and final remarks

The GEV distribution can be applied to the sensory analysis of specialty coffees, whose sensorial panel presents an heterogeneity among consumers.

The probabilities obtained by this distribution show that the sensory analysis of specialty coffees performed by untrained consumers indicates that they are able to differentiate specialty coffees and provide similar scores to the sensory analysis performed by consumers with prior training.

The proposed inference made it possible to attribute some degree of uncertainty regarding the occurrence of sensory scores in the different types of specialty coffees studied and to indicate which group each coffee belongs to with high probability according to the Specialty Grade.

It can be recommended that more intensive training with tasters or the application of the proposed methodology with tasters with international certification should be considered with a view to assessing specialty coffees against a reference score of 9 points, since for the present study, only coffee D has a high probability of presenting this note. It should be noted that according to the analysis protocol provided by Specialty Coffee Association of America, the results of the sensory evaluation vary according to a scale where the grades upper to 9 correspond to exceptional coffee.

The study has some limitations that provide directions for future research, although the GEV distribution is specific for analyzing maximum values, the data generating mechanism truncates the maximum score at 10. This characteristic could be taken into account, fitting the model to truncated data. Some proposals have appeared in the literature to consider truncation in the estimation process by maximum likelihood, but there is no consolidated methodology yet. Therefore, it is a possibility for further studies that may be the subject of future research.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Author details

Gilberto Rodrigues Liska[1*], Luiz Alberto Beijo[2], Marcelo Ângelo Cirillo[3], Flávio Meira Borém[4] and Fortunato Silva de Menezes[5]

1 Department of Agroindustrial Technology and Rural Socioeconomics, Federal University of São Carlos, Araras, São Paulo State, Brazil

2 Department of Statistics, Federal University of Alfenas, Alfenas, Minas Gerais State, Brazil

3 Department of Statistics, Federal University of Lavras, Lavras, Minas Gerais State, Brazil

4 Department of Agricultural Engineering, Federal University of Lavras, Lavras, Minas Gerais State, Brazil

5 Department of Physics, Federal University of Lavras, Lavras, Minas Gerais State, Brazil

*Address all correspondence to: gilbertoliska@ufscar.br

IntechOpen

References

[1] de Oliveira Fassio L, Malta M, Carvalho G, Liska G, de Lima P, Pimenta C. Sensory Description of Cultivars (*Coffea Arabica* L.) Resistant to Rust and Its Correlation with Caffeine, Trigonelline, and Chlorogenic Acid Compounds. Beverages. 2016 Jan 18;2(1):1.

[2] Specialty Coffee Association of America. SCAA Protocols - Cupping Specialty Coffee [Internet]. 2015. Available from: www.scaa.org

[3] Alves HMRA, Volpato MML, Vieira TGC, Borém FM, Barbosa JN. Características ambientais e qualidade da bebida dos cafés do estado de Minas Gerais. Inf Agropecuário. 2011;32(261): 1–12.

[4] Lingle T. The coffee cupper's handbook : systematic guide to the sensory evaluation of coffee's flavor. Fourth edi. Long Beach California: Specialty Coffee Association of America; 2011.

[5] Liska GR, De Menezes FS, Cirillo MA, Borém FM, Cortez RM, Ribeiro DE. Evaluation of sensory panels of consumers of specialty coffee beverages using the boosting method in discriminant analysis. Semin Agrar. 2015;36(6).

[6] Ferreira HA, Liska GR, Cirillo MA, Borém FM, Ribeiro DE, Cortez RM, et al. Selecting A Probabilistic Model Applied to the Sensory Analysis of Specialty Coffees Performed with Consumer. IEEE Lat Am Trans. 2016;14(3).

[7] Malta MR, Chagas SJ de R. Avaliação de compostos não-voláteis em diferentes cultivares de cafeeiro produzidas na região sul de Minas Gerais. Acta Sci - Agron [Internet]. 2009 [cited 2020 Oct 23];31(1):57–61. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=

S1807-86212009000100010&lng=en&nrm=iso&tlng=pt

[8] Chagas E do N, Morais AR de, Cirillo MA, Figueiredo LP, Borém FM. Selection of robust estimator usedin analysis of sensory characteristics and identification of environments conducive to specialty coffee production. Adv Crop Sci [Internet]. 2013 [cited 2020 Oct 23];3(8):515–24. Available from: http://repositorio.ufla.br/jspui/handle/1/13024

[9] Silva FLF, Nascimento GO, Lopes GS, Matos WO, Cunha RL, Malta MR, et al. The concentration of polyphenolic compounds and trace elements in the *Coffea arabica* leaves: Potential chemometric pattern recognition of coffee leaf rust resistance. Food Res Int [Internet]. 2020 Aug;134: 109221. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0963996920302465

[10] Malta MR, Fassio L de O, Liska GR, Carvalho GR, Pereira AA, Botelho CE, et al. Discrimination of genotypes coffee by chemical composition of the beans: Potential markers in natural coffees. Food Res Int [Internet]. 2020 Aug;134: 109219. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0963996920302441

[11] Fassio L de O, Malta MR, Liska GR, Carvalho GR, Botelho CE, Pereira AA, et al. Performance of arabica coffee accessions from the active germplasm bank of Minas Gerais, Brazil as a function of dry and wet processing: a sensory approach. Aust J Crop Sci [Internet]. 2020 Jun 20;14(6):1011–8. Available from: https://www.cropj.com/fassio_14_6_2020_1011_1018.pdf

[12] Fassio LO, Malta MR, Carvalho GR, Pereira AA, Silva AD, Liska GR, et al. Discrimination of Genealogical Groups of Arabica Coffee by the Chemical

Composition of the Beans. J Agric Sci. 2019 Sep 30;11(16):141.

[13] Figueiredo LP, Borém FM, Cirillo MÂ, Ribeiro FC, Giomo GS, Salva TDJG. The Potential for High Quality Bourbon Coffees From Different Environments. J Agric Sci [Internet]. 2013 Sep 15 [cited 2020 Oct 23];5(10):87–98. Available from: http://www.ccsenet.org/journal/index.php/jas/article/view/27842

[14] Barbosa JN, Borem FM, Cirillo MA, Malta MR, Alvarenga AA, Alves HMR. Coffee Quality and Its Interactions with Environmental Factors in Minas Gerais, Brazil. J Agric Sci [Internet]. 2012 Mar 31 [cited 2020 Oct 23];4(5):181–90. Available from: http://www.ccsenet.org/journal/index.php/jas/article/view/13784

[15] Borém FM, Cirillo M, de Carvalho Alves AP, dos Santos CM, Liska GR, Ramos MF, et al. Coffee sensory quality study based on spatial distribution in the Mantiqueira mountain region of Brazil. J Sens Stud. 2020 Apr 1;35(2).

[16] Efron B, Tibshirani RJ. An Introduction to the Bootstrap [Internet]. CRC Press; 1994 [cited 2020 Oct 23]. 456 p. Available from: https://books.google.com.br/books/about/An_Introduction_to_the_Bootstrap.html?id=gLlpIUxRntoC&redir_esc=y

[17] Casella G, Berger RR. Statistical Inference. 2nd ed. Thomson Learning; 2001. 688 p.

[18] Mendes BV de M. Introdução à análise de eventos extremos. Rio de Janeiro: E-papers Serviços Editoriais Ltda; 2004. 232 p.

[19] Coles S. An Introduction to Statistical Modeling of Extreme Values [Internet]. London: Springer London; 2001. 221 p. (Springer Series in Statistics). Available from: http://link.springer.com/10.1007/978-1-4471-3675-0

[20] Blain GC. Dry months in the agricultural region of Ribeirão Preto, state of São Paulo-Brazil: an study based on the extreme value theory. Eng Agrícola [Internet]. 2014 Oct;34(5):992–1000. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-69162014000500018&lng=en&tlng=en

[21] Moral RA, Hinde J, Demétrio CGB. Half-Normal Plots and Overdispersed Models in R : The hnp Package. J Stat Softw [Internet]. 2017;81(10). Available from: https://www.jstatsoft.org/v081/i10

[22] Hartmann M, Moala FA, Mendonça MA. Estudo das precipitações máximas anuais em Presidente Prudente. Rev Bras Meteorol [Internet]. 2011 Dec;26(4):561–8. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-77862011000400006&lng=pt&tlng=pt

[23] Ferreira RV de C, Liska GR. Análise probabilística da temperatura máxima em Uruguaiana, RS. Rev Bras Agric Irrig. 2019 Jul 25;13(3):3390–401.

[24] Ljung GM, Box GEP. On a Measure of Lack of Fit in Time Series Models. Biometrika [Internet]. 1978 Aug;65(2):297. Available from: https://www.jstor.org/stable/2335207?origin=crossref

[25] Martins ALA, Liska GR, Beijo LA, Menezes FS de, Cirillo MÂ. Generalized Pareto distribution applied to the analysis of maximum rainfall events in Uruguaiana, RS, Brazil. SN Appl Sci [Internet]. 2020 Sep 5;2(9):1479. Available from: http://link.springer.com/10.1007/s42452-020-03199-8

[26] Rizzo ML. Statistical Computing with R [Internet]. Chapman and Hall/CRC; 2007. 416 p. Available from: https://www.crcpress.com/Statistical-Computing-with-R/Rizzo/p/book/9781584885450

[27] R Core Team. R: A language and environment for statistical computing. Vienna, Austria; 2018.

[28] RStudio Team. RStudio: Integrated Development for R. [Internet]. Boston: RStudio, Inc; 2015. Available from: http://www.rstudio.com/

[29] Stephenson AG. evd: Extreme Value Distributions. R News [Internet]. 2002;2 (2):31–2. Available from: https://cran.r-project.org/doc/Rnews/Rnews_2002-2.pdf

[30] Crutcher HL. A Note on the Possible Misuse of the Kolmogorov-Smirnov Test. J Appl Meteorol [Internet]. 1975 Dec 1 [cited 2020 Oct 23];14(8):1600–3. Available from: http://journals.ametsoc.org/jamc/article-pdf/14/8/1600/4968526/1520-0450

[31] Bautista EAL, Zocchi SS, Angelocci LR. Fitting the generalized extreme value distribution (GEV) to the maximum wind speed data in Piracicaba, São Paulo, Brazil. Rev Matemática e Estatística. 2004;22(1):95–111.

[32] Tan ML, Samat N, Chan NW, Lee AJ, Li C. Analysis of Precipitation and Temperature Extremes over the Muda River Basin, Malaysia. Water [Internet]. 2019 Feb 6;11(2):283. Available from: http://www.mdpi.com/2073-4441/11/2/283

[33] Sá EAS, Moura CN de, Padilha VL, Campos CGC. Trends in daily precipitation in highlands region of Santa Catarina, southern Brazil. Ambient e Agua - An Interdiscip J Appl Sci [Internet]. 2018 Feb 16;13(1):1–13. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1980-993X2018000100312&lng=en&nrm=iso&tlng=en

[34] Salviano MF, Groppo JD, Pellegrino GQ. Análise de Tendências em Dados de Precipitação e Temperatura no Brasil. Rev Bras Meteorol [Internet]. 2016 Mar;31(1):64–73. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-77862016000100064&lng=pt&tlng=pt

[35] Aguirre AFL, Nogueira DA, Beijo LA. Análise da temperatura máxima de Piracicaba (SP) via distribuição GEV não estacionária: uma abordagem bayesiana. Rev Bras Climatol [Internet]. 2020 Sep 21 [cited 2020 Nov 25];27:496–517. Available from: http://dx.doi.org/10.5380/abclima.v27i0.73763

[36] Liska GR, Sáfadi T, Bortolini J, Beijo LA. Estimativas de velocidade máxima de vento em Piracicaba-SP via Séries Temporais e Teoria de Valores Extremos. Rev Bras Biometria. 2013;31 (2):295–309.

[37] DiCiccio TJ, Efron B. Bootstrap confidence intervals. Stat Sci [Internet]. 1996 [cited 2020 Nov 12];11(3):189–212. Available from: https://projecteuclid.org/euclid.ss/1032280214

[38] Jung K, Lee J, Gupta V, Cho G. Comparison of Bootstrap Confidence Interval Methods for GSCA Using a Monte Carlo Simulation. Front Psychol. 2019 Oct 11;10:1–11.