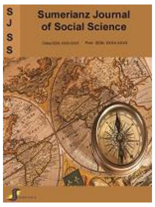
Sumerianz Journal of Social Science, 2024, Vol. 7, No. 4, pp. 57-76
ISSN(e): 2616-8693, ISSN(p): 2617-1716
Website: https://www.sumerianz.com
DOI: https://doi.org/10.47752/sjss.74.57.76



Original Article Open Access

Sampling Distribution and Point and Interval Estimation of Kelley's Percentile Coefficient of Kurtosis

José Moral de la Rubia
School of Psychology, Universidad Autónoma de Nuevo León, Monterrey, Mexico
Email: jose.morald@uanl.edu.mx

Article History

Received: 4 August, 2024

Revised: 25 September, 2024

Accepted: 2 October, 2024

Published: 2 October, 2024

How to Cite

Moral de la Rubia, (2024). Sampling Distribution and Point and Interval Estimation of Kelley's Percentile Coefficient of Kurtosis. *Sumerianz Journal of Social Science,* 7(4): 57-76.

Abstract

There is a robust measure for estimating kurtosis based on quantiles, developed by the American psychologist Truman Lee Kelley, but it is seldom used. This underutilization is likely because it is not available in any statistical package, including R program. Additionally, Kelley reported its standard error when the random variable is drawn from a normal distribution but did not establish its sampling distribution. The objective of this study is to determine its asymptotic sampling distribution and to develop an R script to provide point and interval estimates for this measure. The R program was chosen because it is freely available, developed by the mathematical community, and considered one of the most comprehensive statistical packages currently available. To determine the asymptotic sampling distribution, samples of 100, 500, 1000, 5000, 10000, and 20000 data points were generated from three symmetric distributions: uniform (platykurtic), normal (mesokurtic), and Laplace (leptokurtic). From these 18 source samples, 1000 samples were drawn by resampling with replacement to obtain 18 bootstrap sampling distributions. Normality was then determined using Grubbs (outliers), D'Agostino (symmetry), Anscombe-Glynn (mesokurtosis), and Anderson-Darling, Lilliefors, Shapiro-Francia, and D'Agostino-Belanger-D'Agostino (normality) tests. The script included tests for several assumptions to decide which confidence interval to use: Wald type (normal asymptotic) or bootstrap (Gaussian, percentile, and percentile bias-corrected and accelerated). As an example, the script was applied to a random sample drawn from the raised cosine distribution on waiting time in a social service. It was concluded that the asymptotic sampling distribution of Kelley's kurtosis measure is normal and that this measure appears less specific than classical Pearson and Fisher measures in the presence of a distribution very close to normal. It is suggested to use this script, which may have practical and academic utility.
Keywords: kurtosis; Quantiles; Asymptotic confidence interval; Bootstrap confidence interval; R program.

1. Introduction

Kurtosis is a complex statistical concept that measures aspects of tail heaviness and shoulder thinning, particularly in the case of unimodal distributions (Sharipova, 2024). Its interpretation is simpler with unimodal symmetric distributions (Eberl and Klar, 2024). Kurtosis can be measured using classical methods involving the standardized fourth central moment (Pearson, 1895;1905) and the standardized fourth cumulant (Fisher, 1930), as well as through quantiles (Kelley, 1923;1947). The first two measures consider an axis of symmetry located at the arithmetic mean, while the latter considers the median.

Another robust proposal centered on the median, as an alternative to ' Kelley's, is that of Hogg (1974). This adaptive and robust measure uses partial arithmetic means of the ordered data (lower mean or average of the bottom 5% or 20%, and the bottom 50% of the ascendingly sorted data, as well as the upper mean or average of the top 5% or 20%, and the top 50% of the ascendingly sorted data). A more recent proposal bypasses any measure of central tendency as the axis of symmetry and instead uses a canonical ordering approach with continuous quantitative variables (Eberl and Klar, 2024). Additionally, for qualitative variables, another approach is founded on ordering based on the frequency of nominal categories, taking the mode as the axis of symmetry (Moral, 2023).

In this article, we revisit Kelley (1923) robust approach to measuring kurtosis, which is an important method (Favero and Belfiore, 2019) and is used in applied research, though much less frequently than Fisher's and Pearson's measures. This is primarily because its automated calculation is not available in statistical packages (Chattamvelli and Shanmugam, 2023; Ho and Yu, 2015), including R.

The first objective of this paper is to determine the asymptotic sampling distribution of Kelley's percentile coefficient of kurtosis. Its second objective is to facilitate the use of this coefficient by researchers through a script developed for the R program. Since the mean and variance of this statistic are finite for random variables drawn from a normal distribution and its sampling distribution supports a triangular arrangement, the sufficient conditions for convergence to normality outlined by Nakayama and Tuffin (2024) are met. Therefore, the expectation is that the sampling distribution will be normal.

The objective of the script is to enable point and interval estimation of Kelley's coefficient. The script includes the calculation of a type-Wald or asymptotic interval and bootstrap confidence intervals, with the latter being computed using three approaches: the normal or Gaussian method, the percentile method, and the bias-corrected and accelerated (BCa) percentile method (Canty *et al.*, 2024). The script tests the assumptions of randomness and normality in the random sample of variable X, checks for symmetry and normality in the bootstrap sampling distribution, and calculates the bootstrap estimation bias and jackknife acceleration to decide which approach to use for interval estimation.

The R program is chosen because it is freely available, developed by the mathematical community (Ihaka, 2009), and currently considered one of the best existing statistical programs (Bruce *et al.*, 2020). However, it has the drawback of requiring script writing for its execution, as opposed to the drop-down menu format of other programs (Ramachandran and Tsokos, 2020), such as SPSS (IBM Corporation, 2023).

The following sections of the article present Kelley's percentile coefficient of kurtosis, both centered and not centered at 0. The sampling distribution of the statistic is investigated through bootstrapping from three symmetric distributions with different kurtoses: platykurtic (uniform), mesokurtic (normal), and leptokurtic (Laplace). Methodological details are given in the Method section. A R script is developed for point and interval estimation (Wald or asymptotic and bootstrap types) of the statistic, the script is applied to an example with a variable following a raised arcsine distribution, the results are discussed, and some conclusions are drawn. Additionally, an appendix is provided with an R script for point and interval estimation of Pearson (1895) kurtosis coefficient based on the standardized fourth central moment, Fisher (1930) kurtosis coefficient based on the standardized fourth cumulant, and Hogg (1974) adaptive and robust measure of kurtosis.

2. Kelley's Percentile Coefficient of Kurtosis

In 1923, the American psychologist Truman Lee Kelley (1884-1961) developed a percentile-based measure of kurtosis obtained by the ratio of the Semi-Interquartile Range (*SIQR*) to the Percentile Range (*PR*). It is called the Percentile Coefficient of Kurtosis (PCK). See Equation 1, where $q_p(\mathrm{x})$ represents the *p*-order sample quantile of the random variable x.

$$\mathrm{x} \in \{x_i\}_{i=1}^{n} = \{x_1, x_2, \dots, x_n\} \subseteq \mathrm{X} \quad (1)$$

$$\widehat{PCK}(\mathrm{x}) = \frac{\widehat{SIQR}(\mathrm{x})}{\widehat{PR}(\mathrm{x})} = \frac{\left(q_{0.75}(\mathrm{x}) - q_{0.25}(\mathrm{x})\right)/2}{q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x})} = \frac{q_{0.75}(\mathrm{x}) - q_{0.25}(\mathrm{x})}{2\left(q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x})\right)}$$

Since it is based on quantiles, it is a more robust measure compared to Pearson's measure, which relies on the standardized fourth central moment, and Fisher's measure, which is based on the standardized fourth cumulant. Sampling quantiles can be calculated using rule 8 of the R package, which provides the best inferences when the distribution of the sample or random variable is unknown (Linden, 2023; Sukhoplyuev and Nazarov, 2024).

In R's Rule 8, the order of the sample quantile (p), or the cumulative probability of a datum point in the empirical distribution of n sample data, is defined as the median of the i-th order statistic from random samples of size n drawn from a standard continuous uniform distribution $U[0, 1]$: $Mdn(p) = (\alpha - 1/3) / (\alpha + \beta - 2/3) = (i - 1/3) / (n + 1/3)$, where $p \sim \mathrm{Beta}(\alpha = i, \beta = n + 1 - i)$. Specifically, when probability values are chosen at random, they follow a $U[0, 1]$ distribution; thus, this distribution is used as a non-informative prior when estimating probabilities in Bayesian inference. When i is cleared, the order p of the quantile q_p among the n sample data is obtained. If i is an integer, q_p corresponds to a sample data ($q_p = x_i$). If i is a decimal number, q_p is obtained through linear interpolation. Refer to Equation 2.

There is also the option of using the arithmetic mean (R's rule 6) or the mode (R's rule 7) of the i-th order statistic from random samples of size n drawn from a $U[0, 1]$ distribution. These are among the most commonly used rules, as well as rule 9 for the normal distribution (R Core Team and contributors worldwide, 2024). See Equation 2, in which the n sample data are sorted in ascending order.

$$x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} \quad (2)$$

$$i = \lfloor i \rfloor + (i - \lfloor i \rfloor) = \begin{cases} p(n+1) & Rule\ 6 \\ 1 + p(n-1) & Rule\ 7 \\ 1/3 + p(n+1/3) & Rule\ 8 \\ 3/8 + p(n+1/4) & Rule\ 9 \end{cases}$$

$$x_{(\lfloor i \rfloor)} \le q_p(\mathrm{x}) \le x_{(\lfloor i \rfloor + 1)}$$

$$q_p(\mathrm{x}) = x_{(\lfloor i \rfloor)} + (i - \lfloor i \rfloor)\left(x_{(\lfloor i \rfloor + 1)} - x_{(\lfloor i \rfloor)}\right)$$

The range of *PCK* statistic varies from a minimum of 0 to a maximum of 0.5. When the 1th quartile and the 3th quartile have the same value: $q_{0.75}(\mathrm{x}) = q_{0.25}(\mathrm{x})$, the value of the *PCK* coefficient reaches its minimum of 0. See Equation 3.

$$\widehat{PCK}(\mathrm{x}) = \frac{\widehat{IQR}(\mathrm{x})}{\widehat{PR}(\mathrm{x})} = \frac{q_{0.75}(\mathrm{x}) - q_{0.25}(\mathrm{x})}{2(q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x}))} = \frac{0}{2(q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x}))} = 0 \quad (3)$$

When the distance between the 90th percentile and the 10th percentile is equal to the distance between the third quartile and the first quartile: $q_{0.90}(\mathrm{x}) - q_{0.10}(\mathrm{x}) = q_{0.75}(\mathrm{x}) - q_{0.25}(\mathrm{x})$, the value of the *PCK* coefficient reaches its maximum of 0.5. See Equation 4.

$$\widehat{PCK}(\mathrm{x}) = \frac{\widehat{IQR}(\mathrm{x})}{\widehat{PR}(\mathrm{x})} = \frac{q_{0.75}(\mathrm{x}) - q_{0.25}(\mathrm{x})}{2(q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x}))} = \frac{q_{0.75}(\mathrm{x}) - q_{0.25}(\mathrm{x})}{2(q_{0.75}(\mathrm{x}) - q_{0.25}(\mathrm{x}))} = \frac{1}{2} \quad (4)$$

If Kelley's percentile coefficient of kurtosis is calculated for the standard normal distribution, its value is 0.263. See Equation 5.

$$X \sim N(0, 1) \quad (5)$$

$$PCK(\mathrm{X}) = \frac{IQR(\mathrm{X})}{PR(\mathrm{X})} = \frac{Q_{0.75}(\mathrm{X}) - Q_{0.25}(\mathrm{X})}{2(Q_{0.9}(\mathrm{X}) - Q_{0.1}(\mathrm{X}))}$$

$$= \frac{0.674 - (-0.674)}{2(1.282 - (-1.282))} = \frac{1.348}{2 \times 2.564} = \frac{1.348}{5.128} = 0.263$$

When the normal distribution is used as a benchmark for mesokurtosis, a *PCK* value of 0.263 is interpreted as indicating mesokurtosis, which signifies medium tails. A value lower than 0.263 indicates leptokurtosis, indicating elongated tails, while a value higher than 0.263 indicates platykurtosis, indicating shortened tails.

This interpretation is particularly valid when there is symmetry and a single mode, as seen in the normal distribution. In cases of skewness, it can measure the elongation or contraction of a tail. It is worth noting that the direction of interpretation in Kelley (1923) kurtosis index differs from that of kurtosis metrics derived from standardized central moments or cumulants, as is also the case with Kelley's skewness indices.

For a sample of size n drawn from a variable with a standard normal distribution $N(0, 1)$, the standard deviation of the percentile kurtosis coefficient is given by $\sigma(PCK) = 0.27779 / \sqrt{n}$.

When variable X follows a normal distribution, the sampling distribution of the *PCK*(x) statistic approaches a Gaussian distribution with a mathematical expectation or arithmetic mean $E[PCK(\mathrm{x})] = 0.263$ and standard error $SE[PCK(\mathrm{x})] = 0.27779 / \sqrt{n}$. Therefore, its population value can be estimated by an interval with a confidence level of $(1 - \alpha) \times 100\%$, using Equation 6, where $z_{1-\alpha/2}$ represents the $(1-\alpha/2)$-th quantile of a $N(0, 1)$ distribution.

$$P\left(PCK(\mathrm{X}) \in \left[\widehat{PCK}(\mathrm{x}) - z_{1-\frac{\alpha}{2}}\frac{0.27779}{\sqrt{n}}, \widehat{PCK}(\mathrm{x}) + z_{1-\frac{\alpha}{2}}\frac{0.27779}{\sqrt{n}}\right]\right) = 1 - \alpha \quad (6)$$

Assuming that the sampling distribution of *PCK* converges to a normal distribution with the previously mentioned parameters when n random variables are drawn from a distribution with finite moments and n is large (tending to infinity), Equation 6 provides an asymptotic or Wald-type confidence interval (Wald, 1939; Zepeda-Tello *et al.*, 2022) for *PCK*.

If 0.263 is included in the interval, it indicates mesokurtosis or medium tails. If the upper limit is less than 0.263, it indicates platykurtosis or shortened tails. If the lower limit is greater than 0.263, it indicates leptokurtosis or elongated tails. In both cases, the distribution would deviate from normality. Thus, its standardized form (Equation

7) and its asymptotic interval estimate (Equation 6) can be used as a partial test of normality. This test is partial because normality implies symmetry and a bell-shaped profile in the histogram, along with a characteristic distribution pattern around the mean.

$$Z[\widehat{PCK}(\text{x})] = \frac{(\widehat{PCK}(\text{x}) - 0.26315) \times \sqrt{n}}{0.27779} \sim N(0, 1) \quad (7)$$

Another more general option for calculating the confidence interval is to perform bootstrap sampling by drawing 1000 or more samples with replacement from the original sample, and applying the BCa percentile method. The percentile method is particularly suitable when the bias in the bootstrap estimation is minimal ($|bias| \leq 0.05$), there is symmetry in the bootstrap sampling distribution (tested using a statistical test such as D'Agostino (1970), and the acceleration is minimal ($|a| \leq 0.025$). Alternatively, the normal or Gaussian method can be applied if the bootstrap sampling distribution approximates a normal distribution, resulting in a symmetric interval, unlike the other two methods which produce asymmetric intervals (Caeiro and Mateus, 2024; Efron and Narasimhan, 2020).

3. Percentile Coefficient of Kurtosis Centered at 0.

Pearson (1905), proposed using the normal distribution as a reference for mesokurtosis and defined a measure of kurtosis centered at 0 for the mesokurtic condition by subtracting 3 from the coefficient β_2. This measure is commonly referred to as excess kurtosis. It should be recalled that the coefficient β_2 represents the kurtosis computed using the standardized fourth central moment, and when calculated within the normal distribution its value is 3.

Applying the *PCK* formula to a normal distribution yields 0.263 (Equation 5). Following Pearson's approach, if the normal distribution serves as a reference for mesokurtosis, a kurtosis measure based on percentiles centered at 0 can be derived by subtracting the expected normal value (0.263) from *PCK*. This statistic can be referred to as the percentile coefficient of kurtosis centered at 0 and denoted as *PCKc*. See Equation 8.

$$\widehat{PCK_c}(\text{x}) = \widehat{PCK}(\text{x}) - 0.263 = \frac{q_{0.75}(\text{x}) - q_{0.25}(\text{x})}{2(q_{0.9}(\text{x}) - q_{0.1}(\text{x}))} - 0.263 \quad (8)$$

The range of *PCKc* varies from -0.263 to 0.237. A value of 0 indicates mesokurtosis, a negative value indicates leptokurtosis or thickened tails, and a positive value indicates platykurtosis or thinned tails. This implies an interpretation opposite to that of kurtosis measures based on central moments or cumulants.

If a sample of size n is drawn from a variable with a normal distribution, the standard deviation or error of the corrected percentile kurtosis coefficient is $\sigma(PCK) = 0.27779 / \sqrt{n}$ and its expected value is 0.

Assuming that when n random samples are drawn of a distribution with finite moment and n is large, the sampling distribution of *PCKc* converges to a normal distribution with an expected value or arithmetic mean $\mu(PCKc) = 0$ and standard deviation or error $\sigma(PCKc) = 0.27779 / \sqrt{n}$. Therefore, the population value of *PCKc* can be estimated by an asymptotic interval with a confidence level of $(1 - \alpha) \times 100\%$ using Equation 9.

$$P\left(PCKc(\text{X}) \in \left[\widehat{PCKc}(\text{x}) - z_{1-\frac{\alpha}{2}}\frac{0.27779}{\sqrt{n}}, \widehat{PCKc}(\text{x}) + z_{1-\frac{\alpha}{2}}\frac{0.27779}{\sqrt{n}}\right]\right) = 1 - \alpha \quad (9)$$

If the interval includes 0, it indicates mesokurtosis or medium tails. If the upper limit is less than 0, it indicates platykurtosis or shortened tails. If the lower limit is greater than 0, it indicates leptokurtosis or elongated tails. In both cases, the distribution would deviate from normality. Therefore, its interval estimation can serve as a partial test of normality, recognizing that normality implies both symmetry and mesokurtosis, resulting in a bell-shaped profile in the histogram.

Another alternative to the asymptotic confidence interval is the bootstrap confidence interval, specifically using the BCa percentile method, which is nonparametric and does not require assumptions about the distribution. As previously noted, when bias and acceleration are minimal, and there is symmetry in the bootstrap sampling distribution, the percentile method is suitable. Alternatively, if the bootstrap sampling distribution approximates normality, the normal or Gaussian method can be applied.

4. Method

In relation to the first objective of determining the sampling distribution of the *PCK* statistic (Kelley, 1923), 18 source samples of six sizes (101, 501, 1001, 5001, 5001, 10001, and 20001 data points) were generated from three symmetric continuous distributions centered at 0. These distributions were: one platykurtic (uniform with threshold parameters: a = -3 and b = 3), one mesokurtic (normal with location parameter μ = 0 and scale parameter σ = 1), and one leptokurtic (Laplace with location parameter μ = 0 and scale parameter β = 1).

The samples were created using inverse transform sampling (Baumgarten and Patel, 2022) from a sequence of cumulative probability values (quantile orders) equally spaced between 0.001 and 0.999. An odd number of data points was chosen to ensure that the modal value of 0 would appear in each source sample. Additionally, equispaced quantile orders (p) were selected so that the quantile values $_{(p)}$ on both sides of the peak, located at 0, would be exactly the same with opposite signs, ensuring the source distributions were perfectly symmetric.

From each of the 18 source samples, 1000 random samples were drawn by sampling with replacement, and in each of these 1000 bootstrap samples, the *PCK* statistic was calculated, obtaining the corresponding bootstrap sampling distribution of *PCK*. In each of these 18 bootstrap sampling distributions, the following properties were tested: outliers (Grubbs, 1969), symmetry (D'Agostino, 1970), kurtosis (Anscombe and Glynn, 1983), and normality (Anderson and Darling, 1952; D'Agostino and Pearson, 1973; D'Agostino *et al.*, 1990; Kolmogorov, 1933; Lilliefors, 1967; Smirnov, 1948); as well as Shapiro and Francia (1972), with standardization by Royston (1993). Each of these tests has different rationales (Demir, 2022).

The Anderson-Darling test is based on the sum of the standardized quadratic distances between the empirical and theoretical distribution functions of the n sample data. The Kolmogorov-Smirnov test relies on the maximum vertical distance between the empirical and theoretical distributions of the n sample data (Smirnov, 1948), with critical values obtained by Monte Carlo simulation (Lilliefors, 1967) instead of using the Kolmogorov's distribution (Kolmogorov, 1933). The Shapiro-Francia test is based on the shared variance between the theoretical and empirical quantiles (Shapiro and Francia, 1972), to which Royston (1993) adds a standardization procedure for better convergence of the test statistic to a $N(0, 1)$ distribution. The D'Agostino-Pearson test of is based on Pearson's coefficients of skewness and kurtosis (D'Agostino and Pearson, 1973). Its statistic is the sum of squares of the standardized values of the skewness and kurtosis measures, which are derived from standardized central moments. D'Agostino *et al.* (1990), improved the convergence of the test statistic to a chi-square distribution possessing two degrees of freedom, using D'Agostino's procedure to standardize the skewness measure and Anscombe-Glynn procedure to standardize the kurtosis measure.

In addition, the asymptotic standard error (Kelley, 1923) was calculated and its difference from the bootstrap standard error (Efron, 2022) was tested using the chi-square test for one variance (Snedecor and Cochran, 1989), specifying the square of the asymptotic standard error as the null hypothesis (Kyoungjae *et al.*, 2022).

To achieve the second objective of estimating the *PCK* coefficient both by interval and point estimation, a script was developed for the R program (Denis, 2020). The script begins by testing the randomness of the sample of variable X using the runs test developed by Wald and Wolfowitz (1943) and visualizing the sample distribution with a histogram overlaid with density and normal curves. The number of class intervals of uniform width is determined using the optimal width rule proposed by Freedman and Diaconis (1981), which works well with a wide variety of distributions except for the arcsine distribution (Chattamvelli and Shanmugam, 2021), for which the square root rule is a preferable option according to a study conducted by Moral (2024).

The density is estimated using the Epanechnikov's kernel function (Epanechnikov, 1969), which provides the best minimization of the integrated mean square error (Sholl and Steckel, 2022). The bandwidth is determined using the method developed by Sheather and Jones (1991), which minimizes the asymptotic integrated mean square error, making it a versatile and effective approach (Jiang *et al.*, 2020).

The script continues with point estimation and the calculation of the asymptotic confidence interval. If the sampling distribution conforms to normality, this type of interval is suitable. The process of estimation concludes with the bootstrap confidence interval (Canty *et al.*, 2024).

The skewness of the bootstrap sampling distribution is tested using D'Agostino's test and normality is assessed with the Shapiro- Francia test. Additionally, the bootstrap estimate bias, which is the difference between the mean PCK in the bootstrap sampling distribution and the PCK estimate in the original sample, is reported, along with the jackknife acceleration, which measures the rate of change in the standard deviation of the estimates as the data change (Efron and Narasimhan, 2022).

This script is applied to a random sample composed of 1000 data points, which was generated from a raised cosine distribution with a location parameter $\mu = 11$, a scale parameter $s = 9$, and a range from 2 to 20 (Chattamvelli and Shanmugam, 2021). To give specific meaning to the continuous quantitative variable, it represents the waiting time (in days) for the resolution of a support request in a social service aimed at adolescent mothers (Ahsanullah *et al.*, 2019; McIntyre and Chow, 2020). Because the raised arcsine function lacks an analytical quantile function (Ahsanullah *et al.*, 2019), the acceptance-rejection procedure was used to obtain the probability density function (Equation 10), ensuring that the generated random samples fit the theoretical distribution well (Chowdhury, 2023).

$$f_X(x|\mu = 11, s = 9) = \frac{1}{2s}\left[1 + cos\left(\frac{x-\mu}{s}\pi\right)\right] = \frac{1}{18}\left[1 + cos\left(\frac{x-11}{9}\pi\right)\right] \quad (10)$$

$$x \in [2, 20]$$

5. Results

5.1. Bootstrap Sampling Distribution of *PCK* Statistic

In the histogram of the 18 bootstrap sampling distributions, the unimodal bell-shaped profile of a normal distribution is clearly visible. Similarly, in the normal quantile-quantile plot, the points show a good alignment around the central 45-degree line. These two plots are shown for the bootstrap sampling distributions of the 20001-data source samples derived from the uniform (Figure 1), normal (Figure 2), and Laplace (Figure 3) distributions, as well as their corresponding source samples (Figure 4).

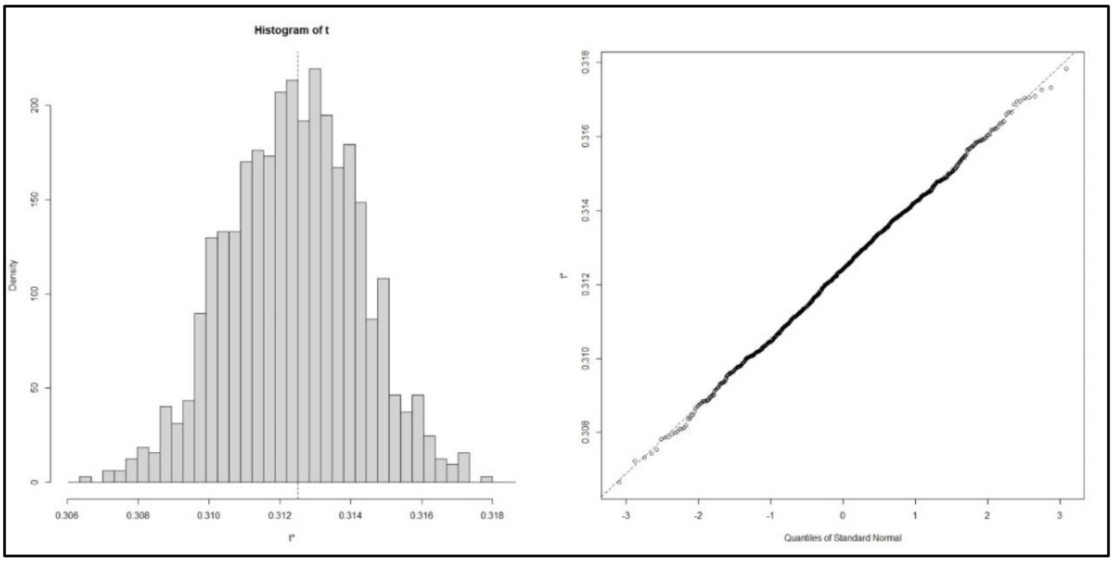


Figure-1. Histogram and normal Q-Q plot of the bootstrap sampling distribution derived from the source sample of 20001 data points following a *U*[-3, 3] distribution.

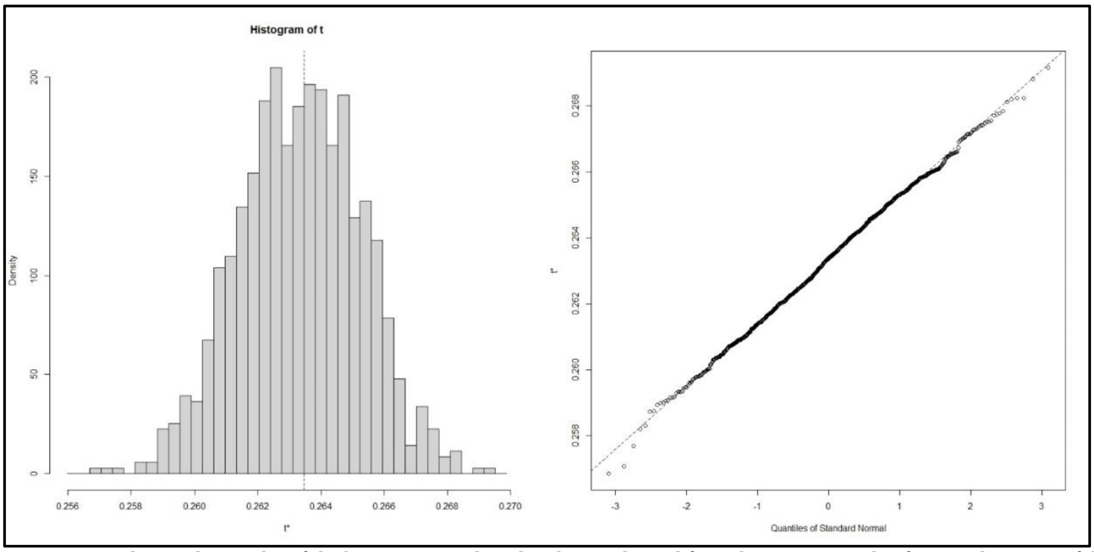


Figure-2. Histogram and normal Q-Q plot of the bootstrap sampling distribution derived from the source sample of 20001 data points following a *N*(0, 1) distribution.

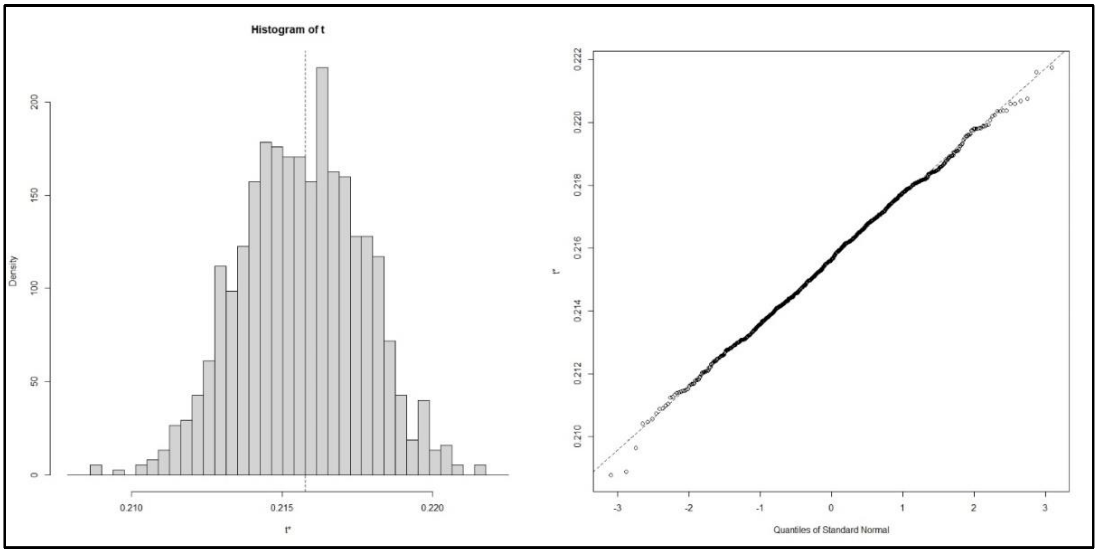


Figure-3. Histogram and normal Q-Q plot of the bootstrap sampling distribution derived from the source sample of 20001 data points following a Laplace($\mu = 0$, $\beta = 1$) distribution.

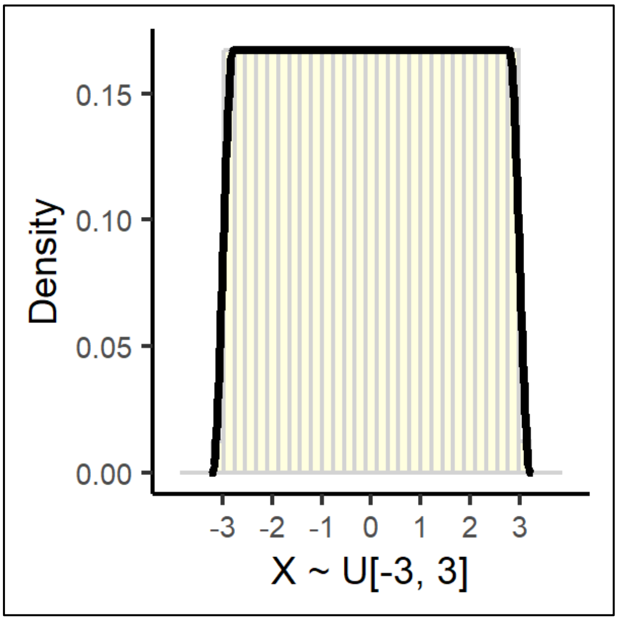


Figure-4.1. Histograms with overlaid density curves for the source samples of 20,001 data points generated from *U*[-3, 3] distribution.

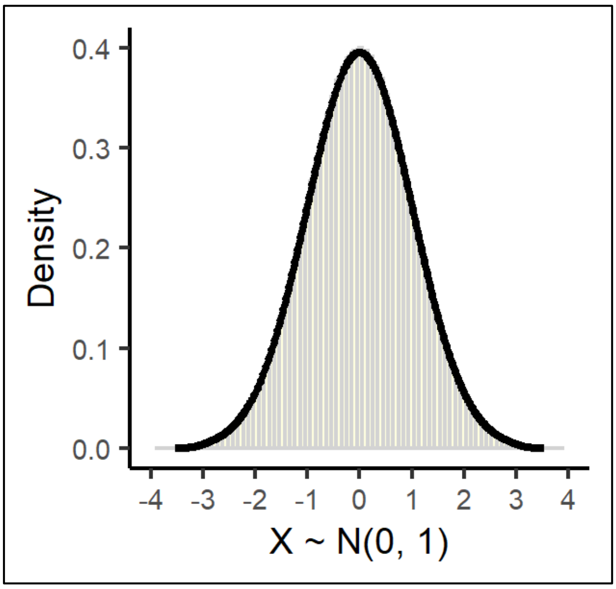


Figure-4.2. Histograms with overlaid density curves for the source samples of 20,001 data points generated from *N*(0, 1), and distribution

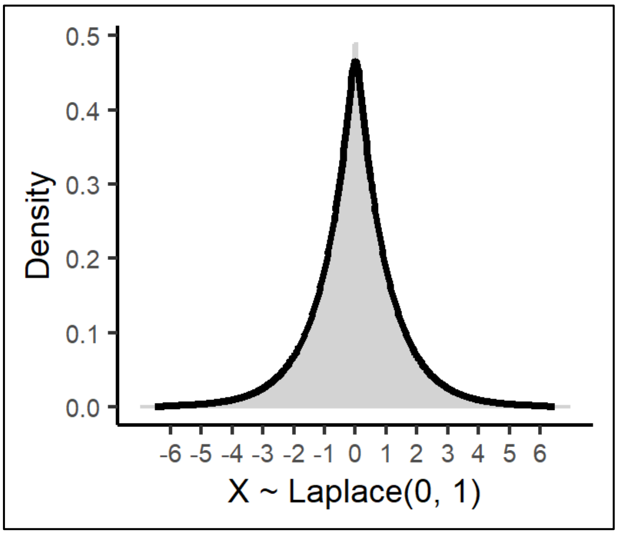


Figure-4.3. Histograms with overlaid density curves for the source samples of 20,001 data points generated from Laplace(μ = 0, β = 1) distribution.

Under the hypothesis of normality or convergence to normality of the sampling distribution of the *PCK* statistic, 15 of the 18 bootstrap sampling distributions of the *PCK* lack outliers according to Grubbs' test at the 5% significance level. The exceptions occur in the bootstrap sampling distributions derived from the 101-data source sample generated from the Laplace distribution and from the 1001-data source samples generated from the normal and Laplace distributions. In all three cases, the outlier is the maximum value, and the null hypothesis would hold at the 1% significance level (Table 1).

In addition, 16 of the 18 bootstrap sampling distributions of the *PCK* statistic are symmetric according to D'Agostino's test at the 5% significance level. Detours are observed in the bootstrap sampling distributions derived from the source samples generated from the Laplace distribution. In both cases, the skewness is right-tailed ($\sqrt{b_1}$ = 0.179 and 0.164, respectively). However, at a 1% significance level, the null hypothesis of symmetry would hold (Table 1).

Additionally, 17 of the 18 bootstrap sampling distributions exhibit mesokurtosis according to the Anscombe-Glynn test at the 5% significance level. The exception is the bootstrap sampling distribution derived from the 1001-data source sample generated from the Laplace distribution, which shows leptokurtosis or heavy tails ($b_2 = 3.375 > 3$). However, if the significance level is lowered to 1%, the null hypothesis of mesokurtosis would hold (Table 1).

Consequently, 17 of the 18 bootstrap sampling distributions fit a normal distribution according to the Anderson-Darling, Kolmogorov-Smirnov-Lilliefors, Shapiro-Francia with Royston's standardization, and D'Agostino-Belanger-D'Agostino tests at the 5% significance level. The exception is the bootstrap sampling distribution derived from the 1001-data source sample generated from the Laplace distribution, which deviates from normality according to the Shapiro-Francia test at the 5% significance level and the D'Agostino-Belanger-D'Agostino test at the 1% significance level (Table 1). However, the null hypothesis of normality is supported for this bootstrap sampling distribution by the Anderson-Darling and Kolmogorov-Smirnov-Lilliefors tests at the 5% significance level.

Table-1. Probability values of the outlier, symmetry, kurtosis, and normality tests for bootstrap sampling distribution of *PCK* statistic

Distribut.	n	G	D'A	AG	AD	L	SF	DBD
Uniform	101	0.8553	0.1935	0.6213	0.5060	0.1729	0.5560	0.3801
U[-3, 3]	501	0.6106	0.2320	0.8489	0.7943	0.9357	0.4912	0.4807
	1001	0.2376	0.7309	0.6897	0.8857	0.7051	0.8681	0.8704
	5001	1	0.5473	0.0561	0.1386	0.2992	0.1610	0.1345
	10001	0.7974	0.8802	0.2706	0.4667	0.2863	0.8007	0.5389
	20001	0.8757	0.5609	0.3034	0.4229	0.5234	0.7849	0.4972
Normal	101	0.2133	0.8553	0.8322	0.9176	0.7063	0.9462	0.9617
N(0, 1)	501	0.3697	0.0595	0.5959	0.4197	0.5484	0.3420	0.1471
	1001	0.0349*	0.1523	0.0975	0.3794	0.2356	0.1045	0.0910
	5001	1	0.4664	0.0755	0.4805	0.7975	0.2614	0.1580
	10001	0.9421	0.8345	0.6603	0.9142	0.9703	0.9519	0.8884
	20001	0.3524	0.5006	0.2913	0.2064	0.3191	0.5206	0.4567
Laplace	101	0.0332*	0.0851	0.7861	0.6120	0.3386	0.2129	0.2189
$\mu = 0$	501	0.2884	0.0222*	0.4825	0.2092	0.2505	0.1506	0.0573
$\beta = 1$	1001	0.0107*	0.0338*	0.0285*	0.1723	0.2181	0.0137*	0.0095**
	5001	1	0.3830	0.0894	0.2621	0.5856	0.1438	0.1615
	10001	0.8397	0.8617	0.7516	0.9086	0.7344	0.9537	0.9368
	20001	0.3343	0.5194	0.3257	0.2132	0.2762	0.4608	0.5013

Note. D = distribution: U = uniform *U*[-3, 3], N = normal *N*(0, 1), L = Laplace(0, 1); Grubbs' outlier test (G); D'Agostino asymmetry test (D'A); Anscombe-Glynn kurtosis test (AG), and normality test: Anderson-Darling (AD), Lilliefors (L), Shapiro-Francia with Royston's standardization (SF), and D'Agostino-Berlanger-D'Agostino (DBD). Significant test with a significance level of: * 5% and ** 1%.

According to the hypothesis of normality or convergence to normality, the asymptotic standard error and bootstrap standard error are statistically equivalent at the 5% significance level for the six sizes of source samples generated from the normal and uniform distributions. The differences between these two errors appear in the source samples generated from the Laplace distribution and become smaller and smaller as the sample size increases, until finally the difference is not significant at the highest sample size of 20001. See Table 2.

Table-2. Asymptotic and bootstrap standard errors for *PCK* statistic

Distribution	*n*	*ase*	*bse*	|*diff*|	chi-2 statistic	*p*-value
Uniform	101	0.0276	0.0276	0.000050	995.422	0.9480
U[-3, 3]	501	0.0124	0.0119	0.000522	916.775	0.0604
	1001	0.0088	0.0087	0.000086	979.501	0.6714
	5001	0.0039	0.0038	0.000095	951.349	0.2853
	10001	0.0028	0.0028	0.000002	1000.741	0.9571
	20001	0.0020	0.0019	0.000079	920.150	0.0722
Normal	101	0.0276	0.0281	0.000430	1030.358	0.4782
N(0, 1)	501	0.0124	0.0124	0.000051	990.796	0.8657
	1001	0.0088	0.0088	0.000051	1010.539	0.7856
	5001	0.0039	0.0039	0.000012	1005.045	0.8809
	10001	0.0028	0.0028	0.000042	1029.151	0.4948
	20001	0.0020	0.0019	0.000048	950.870	0.2803

Laplace	101	0.0276	0.0293	0.001626	1119.972	0.0088**
μ = 0	501	0.0124	0.0131	0.000664	1108.790	0.0170*
β = 1	1001	0.0088	0.0092	0.000458	1106.049	0.0199*
	5001	0.0039	0.0041	0.000216	1111.830	0.0143*
	10001	0.0028	0.0029	0.000167	1122.854	0.0074**
	20001	0.0020	0.0020	0.000058	1059.364	0.1806

Note. *n* = sample size, *ase* = asymptotic standard error, *bse* = bootstrap standard error, |*diff*| = *bse* - *ase* =absolute difference between two errors, chi-2 statistic = testing statistic of Snedecor-Cochran chi-square test for one-variance (*ase* as the null hypothesis), *p*-value = probability value for a two-tailed test. Significant difference with a significance level of: * 5% and ** 1%.

Table 3 illustrates that bias values in the bootstrap estimations were minimal across 18 bootstrap sampling distributions, ranging from 0.00002 to 0.00120, with an average absolute value of 0.00026 (|*bias*| < 0.05). Acceleration values were also notably low, consistently positive ranging from 0.0007 to 0.0145, with an average of 0.0044 (|*a*| < 0.025). It is noteworthy that there was a significant positive correlation between jackknife acceleration and the skewness of bootstrap sampling distributions, assessed by the standardized third central moment coefficient, at a 5% significance level in a two-tailed test (Spearman's correlation: $r_S[18] = 0.521$, $t[16] = 2.443$, $p_{2\text{-}tailed} = 0.027$; $p_{1\text{-}tailed} = 0.013$). These findings are indicative of the bootstrap sampling distributions' good fit to normality.

Table-3. Bootstrap estimation bias, jackknife acceleration, and skewness in the bootstrap sampling distribution of *PCK* statistic.

Distribution	n	Point estimate	Boot estimate	*bias*	*a*	$\sqrt{b_1}$(bsd$t)
Uniform	101	0.3125	0.311303	-0.001197	0.009434	-0.10022
U[-3, 3]	501	0.3125	0.312639	0.000139	0.004288	0.092098
	1001	0.3125	0.312382	-0.000118	0.003037	0.026451
	5001	0.3125	0.312453	-0.000047	0.001360	0.046302
	10001	0.3125	0.312642	0.000142	0.000962	-0.011585
	20001	0.3125	0.312389	-0.000111	0.000680	-0.044737
Normal	101	0.261208	0.260982	-0.000226	0.014477	0.014019
N(0, 1)	501	0.263044	0.263350	0.000306	0.006469	0.147084
	1001	0.26326	0.26323	-0.00003	0.004572	0.110398
	5001	0.263432	0.263416	-0.000016	0.002044	0.056053
	10001	0.263453	0.263613	0.00016	0.001445	-0.016062
	20001	0.263464	0.263338	-0.000126	0.001022	-0.051829
Laplace	101	0.212634	0.213756	0.001122	0.014311	0.132968
μ = 0	501	0.215185	0.215766	0.000581	0.006423	0.179137
β = 1	1001	0.215487	0.21559	0.000103	0.004541	0.164277
	5001	0.215726	0.215742	0.000015	0.002031	0.067156
	10001	0.215756	0.215936	0.00018	0.001436	-0.013397
	20001	0.215771	0.215642	-0.000129	0.001015	-0.049569

Note. *N* = sample size of the source sample, Point estimate = value of *PCK* in the source sample, Boot estimate = mean(bsd$t) = bootstrap estimate or mean of *PCK* in the bootstrap sampling distribution, *bias* = difference between point estimate y bootstrap estimate, *a* = jackknife acceleration or rate of change in the standard deviation of the estimate as the removed data changes, $\sqrt{b_1}$(bsd$t) = skewness of the bootstrap sampling distribution using measure based on the standardized third central moment.

Table 4 presents interpretive guidelines for the three distributions under study. Given the bootstrap sampling distributions' fit to the normal distribution, including the one generated from the 1001-data source sample derived from the Laplace distribution according to two tests (Anderson-Darling and Kolmogorov-Smirnov-Lilliefors), both asymptotic confidence intervals and bootstrap normal confidence intervals are shown at 90%, 95%, and 99% confidence levels.

Table-4. Point estimate and bootstrap normal confidence intervals for PCK statistic

Distribution	n	Me-thod	Point estimate	Bootstrap confidence interval		
				90%	95%	99%
Uniform	101	a	0.3125	(0.2670, 0.3580)	(0.2583, 0.3667)	(0.2413, 0.3837)
U[-3, 3]		b		(0.2683, 0.3591)	(0.2596, 0.3678)	(0.2426, 0.3848)
	501	a	0.3125	(0.2921, 0.3329)	(0.2882, 0.3368)	(0.2805, 0.3445)
		b		(0.2928, 0.3319)	(0.2891, 0.3357)	(0.2817, 0.3430)
	1001	a	0.3125	(0.2981, 0.3269)	(0.2953, 0.3297)	(0.2899, 0.3351)
		b		(0.2983, 0.3269)	(0.2956, 0.3297)	(0.2902, 0.3350)
	5001	a	0.3125	(0.3060, 0.3190)	(0.3048, 0.3202)	(0.3024, 0.3226)
		b		(0.3062, 0.3189)	(0.3050, 0.3201)	(0.3027, 0.3224)
	10001	a	0.3125	(0.3079, 0.3171)	(0.3071, 0.3179)	(0.3053, 0.3197)
		b		(0.3078, 0.3169)	(0.3069, 0.3178)	(0.3052, 0.3195)
	20001	a	0.3125	(0.3093, 0.3157)	(0.3087, 0.3163)	(0.3074, 0.3176)
				(0.3096, 0.3156)	(0.3090, 0.3162)	(0.3079, 0.3173)

Normal	101	a	0.2612	(0.2157, 0.3067)	(0.2070, 0.3154)	(0.1900, 0.3324)
N(0, 1)		b		(0.2153, 0.3076)	(0.2064, 0.3165)	(0.1891, 0.3337)
	501	a	0.2630	(0.2426, 0.2835)	(0.2387, 0.2874)	(0.2311, 0.2950)
		b		(0.2424, 0.2831)	(0.2385, 0.2870)	(0.2309, 0.2946)
	1001	a	0.2633	(0.2488, 0.2777)	(0.2461, 0.2805)	(0.2406, 0.2859)
		b		(0.2488, 0.2778)	(0.2460, 0.2806)	(0.2405, 0.2860)
	5001	a	0.2634	(0.2570, 0.2699)	(0.2557, 0.2711)	(0.2533, 0.2736)
		b		(0.2570, 0.2699)	(0.2557, 0.2712)	(0.2533, 0.2736)
	10001	a	0.2635	(0.2589, 0.2680)	(0.2580, 0.2689)	(0.2563, 0.2706)
		b		(0.2587, 0.2679)	(0.2578, 0.2688)	(0.2560, 0.2706)
	20001	a	0.2635	(0.2602, 0.2667)	(0.2596, 0.2673)	(0.2584, 0.2685)
		b		(0.2604, 0.2667)	(0.2598, 0.2673)	(0.2587, 0.2685)
Laplace	101	a	0.2126	(0.1672, 0.2581)	(0.1585, 0.2668)	(0.1414, 0.2838)
μ = 0		b		(0.1634, 0.2597)	(0.1542, 0.2689)	(0.1361, 0.2869)
β = 1	501	a	0.2152	(0.1948, 0.2356)	(0.1909, 0.2395)	(0.1832, 0.2472)
		b		(0.1931, 0.2361)	(0.1890, 0.2402)	(0.1809, 0.2483)
	1001	a	0.2155	(0.2010, 0.2299)	(0.1983, 0.2327)	(0.1929, 0.2381)
		b		(0.2002, 0.2306)	(0.1973, 0.2335)	(0.1916, 0.2392)
	5001	a	0.2157	(0.2093, 0.2222)	(0.2080, 0.2234)	(0.2056, 0.2258)
		b		(0.2089, 0.2225)	(0.2076, 0.2238)	(0.2050, 0.2264)
	10001	a	0.2158	(0.2112, 0.2203)	(0.2103, 0.2212)	(0.2086, 0.2229)
		b		(0.2107, 0.2204)	(0.2098, 0.2213)	(0.2080, 0.2232)
	20001	a	0.2158	(0.2125, 0.2190)	(0.2119, 0.2196)	(0.2107, 0.2208)
		b		(0.2126, 0.2192)	(0.2119, 0.2199)	(0.2107, 0.2211)

Note. *n* = sample size. Method: a = asymptotic and b = bootstrap using normal method.

5.2. Script for R

The script for the point and interval estimation of Kelley's percentile coefficient of kurtosis (*PCK*) and its variant centered at 0 (*PCKc*), analogous to Pearson's concept of excess kurtosis, is presented below. The 95% asymptotic and bootstrap confidence intervals are computed, with the latter calculated using three methods: normal, percentile, and BCa percentile (Canty *et al.*, 2024; Efron and Narasimhan, 2020). Sample quantiles are calculated using R rule 8 (R Core Team and contributors worldwide, 2024).

If the sample x of the quantitative variable X ($x \subset X$) is random, with a size of at least 30 participants, and follows a normal distribution, the best choice is the asymptotic interval. To ensure this, randomness is tested using the Wald-Wolfowitz runs test. For the distributional assumption, skewness is assessed using D'Agostino's test, kurtosis with the Anscombe-Glynn test, and normality through three different tests: the Anderson and Darling (1952) test, the Shapiro and Francia (1972) test with Royston (1993) standardization, and the test by D'Agostino *et al.* (1990). Additionally, a histogram with overlaid density and normal curves is included. The density is estimated using Epanechnikov's kernel, and the bandwidth is determined by the Sheather-Jones method, which minimizes the asymptotic mean integrated square error.

In case the sampling distribution of X does not conform to normality and the bootstrap standard error is not large (*bse* = *sd*(bsd$t) < *R*(x) / 4 = [*max*(x) - *min*(x)] / 4), the best choice is the bootstrap interval. If the bootstrap sampling distribution of the *PCK* statistic (bsd$t) is normal, the normal method is preferred. If the bootstrap sampling distribution is not normal, but the bias and acceleration are small ($|bias| \leq 0.05$ and $|a| \leq 0.025$) and symmetry is present (D'Agostino, 1970), the percentile method is a good choice. In the case of non-normality, skewness, and non-small bias and acceleration, the BCa percentile method is recommended.

The script includes computing the probability of mesokurtosis (H_0: PCKc = 0) using the bootstrap procedure in both a one-tailed and two-tailed test. If the two-tailed bootstrap probability exceeds the significance level (alpha of 0.05), the null hypothesis is retained; otherwise, it is rejected.

Running this script requires loading six R libraries: randtests (Caeiro and Mateus, 2024), ggplot2 (Wickham *et al.*, 2024), outliers (Komsta, 2022), moments (Komsta and Novomestky, 2022), nortest (Gross and Ligges, 2022), and boot (Canty *et al.*, 2024). It can be run on a personal computer using the R program (R Core Team, 2020) or RStudio (R Studio Team, 2020), which allows the graphics to be saved as PDF or image files (high-definition JPEG). Another option is to run it online from the Snippets - Run R code online site at the address: https://rdrr.io/snippets/

With a small sample size (20 to 29), the significance level can be increased to 0.1, and a 90% confidence interval can be used (Lakens, 2022). Results are displayed rounded to 4 decimal places, except for acceleration, which is rounded to 6 decimal places. Aspects of the script that can be modified to suit the researcher's needs, including the vector of scores, are indicated in blue.

```
# Vector of scores. Put numerical scores or records separated by commas.
x <- c(13.551419, 15.258590, …, 5.912899)
# Load libraries
library(randtests)
```

```
library(outliers)
library(moments)
library(nortest)
library(boot)
library(ggplot2)
# Randomness for x
runse <- runs.test(x, alternative = "two.sided", threshold = median(x), pvalue = 'exact')
runsa <- runs.test(x, alternative = "two.sided", threshold = median(x), pvalue = 'normal')
cat("Wald-Wolfowitz runs test. Criterion: median", "\n")
cat("Number of runs: r =", runse$runs, "\n")
cat("n_0 = #(x_i < mdn(x)) =", runse$parameter["n1"],"and", "n_1 = #(x_i > mdn(x)) =", runse$parameter["n2"], "\n")
cat("n = n_0 + n_1 =", runse$parameter["n"], "\n")
cat("Two-tailed exact probability value: p =", round(runse$p.value, 4), "\n")
cat("Mean: M(R|n_0, n_1) =", runse$mu, "and", "Standard deviation: DE(R|n_0, n_1) =", round(sqrt(runse$var), 4), "\n")
cat("Standardized number of runs: z_r =", round(runse$statistic, 4), "\n")
cat("Two-tailed asymptotic probability value: p-value =", round(runsa$p.value, 4), "\n")
# Normality for x
grubbs.test(x, type = 10, opposite = FALSE, two.sided = FALSE)
agostino.test(x)
anscombe.test(x)
ad.test(x)
lillie.test (x)
sf.test(x)
agostino <- agostino.test(x)
ag <- anscombe.test(x)
K2 <- (agostino$statistic["z"]^2 + ag$statistic["z"]^2)
p <- pchisq(K2, df = 2, lower.tail = FALSE)
cat("D'Agostino-Berlanger-D'Agostino Normality test: x", "\n")
cat("k^2(x) =", round(K2, 4),",","p_value =", round(p, 4), "\n")
# Point estimate of PCK
n <- length(x)
cat("Sample size: n =",n,"\n")
q0.1 <- quantile(x, 0.10, type = 8)
q0.25 <- quantile(x, 0.25, type = 8)
q0.75 <- quantile(x, 0.75, type = 8)
q0.9 <- quantile(x, 0.90, type = 8)
PCK <- (q0.75 - q0.25) / (2*(q0.9 - q0.1))
cat("Kelley's Percentile Coefficient of Kurtosis: PCK = ",PCK,"\n")
# Histogram of x with overlaid density and normal curves
dens <- density(x, kernel = "epanechnikov", bw = "SJ")
fd <- 2 * (q0.75 - q0.25) / n^(1/3)
plot <- ggplot(data = data.frame(x = x), aes(x = x)) +
geom_histogram(binwidth = fd, aes(y = after_stat(density)), fill = "lightyellow", color = "black") + geom_line(data = data.frame(x = dens$x, y = dens$y), aes(x = x, y = y), color = "black", linewidth = 1) + stat_function(fun = dnorm, args = list(mean = mean(x), sd = sd(x)), color = "red", linewidth = 0.5) + labs(x = "Values of X", y = "Density") +
theme(axis.text.x.bottom = element_text(size = 7), axis.text.y = element_text(size = 7), axis.title.x = element_text(size = 9), axis.title.y = element_text(size = 9), panel.background = element_rect(fill = "white"), panel.grid.major = element_blank(), panel.grid.minor = element_blank(), axis.line = element_line(color = "black"))
jpeg("histogram.jpeg", width = 800, height = 600, units = "px", res = 300)
print(plot)
dev.off()
# Asymptotic confidence interval for PCK
ase <- 0.27779 / sqrt(length(x))
alpha <- 0.05
LL_PCK = PCK - qnorm(1-alpha/2) * ase
UL_PCK = PCK + qnorm(1-alpha/2) * ase
cat("Asymptotic standard error of PCK and PCKc: ase(PCK) = ase(PCKc) =", round(ase, 4), "\n")
cat("Asymptotic confidence interval", (1 - alpha) * 100, "% for PCK: 95% CI (", round(LL_PCK, 4),",", round(UL_PCK, 4),")\n")
E_PCK_normal <- (qnorm (0.75) - qnorm (0.25)) / (2*(qnorm (0.9) - qnorm (0.1)))
cat("PCK value corresponding to a normal distribution (mesokurtosis) =", round(E_PCK_normal, 4), "\n")
#Bootstrap confidence interval for PCK
set.seed(123)
```

```
bsd_1 <- boot(data = x, function(x, i) {(quantile(x[i], 0.75, type = 8) - quantile(x[i], 0.25, type = 8)) / (2*(quantile(x[i], 0.90, type = 8) - quantile(x[i], 0.10, type = 8)))}, R = 1000)
PCK_jack <- numeric(n)
for (i in 1:n) {x_jack <- x[-i]
PCK_jack[i] <- (quantile(x_jack, 0.75, type = 8) - quantile(x_jack, 0.25, type = 8)) / (2*(quantile(x_jack, 0.90, type = 8) - quantile(x_jack, 0.10, type = 8)))}
accel <- sum((mean(PCK_jack) - PCK_jack)^3) / (6 * sum((mean(PCK_jack) - PCK_jack)^2)^(3/2))
cat("Bootstrap sampling distribution of PCK statistic (bsd_1$t)", "\n")
agostino.test(bsd_1$t, alternative = "two.sided")
sf.test(bsd_1$t)
ad.test(bsd_1$t)
dagostino <- agostino.test(bsd_1$t)
a_g <- anscombe.test(bsd_1$t)
Ksq <- (dagostino$statistic["z"]^2 + a_g$statistic["z"]^2)
p_value <- pchisq(Ksq, df = 2, lower.tail = FALSE)
cat("D'Agostino-Berlanger-D'Agostino Normality test: bsd_1$t", "\n")
cat("k^2(x) =", round(Ksq, 4),",","p_value =", round(p_value, 4), "\n")
t1 <- mean(bsd_1$t)
bsd_1
cat("Bootstrap estimate of PCK: t* =", t1, "\n")
cat("Jackknife acceleration for PCK: a =", round(accel, 6), "\n")
plot(bsd_1)
boot.ci(bsd_1, conf = 0.95, type = c("norm", "bca", "perc"))
# Point estimate of PCKc
PCKc <- PCK - E_PCK_normal
cat("Percentile coefficient of kurtosis centered at 0 = ", PCKc,"\n")
# Asymptotic confidence interval for PCKc
LL_PCKc <- PCKc - qnorm(1-alpha/2) * ase
UL_PCKc <- PCKc + qnorm(1-alpha/2) * ase
cat("Asymptotic confidence interval", (1 - alpha) * 100, "% for PCKc: 95% CI (", round(LL_PCKc, 4),",", round(UL_PCKc, 4),")\n")
cat("PCKc value corresponding to a normal distribution (mesokurtosis) = 0", "\n")
#Bootstrap confidence interval for PCKc
set.seed(123)
bsd_2 <-boot(data = x, function(x, i) {(quantile(x[i], 0.75, type = 8) - quantile(x[i], 0.25, type = 8)) / (2*(quantile(x[i], 0.90, type = 8) - quantile(x[i], 0.10, type = 8))) - E_PCK_normal}, R = 1000)
PCKC_jack <- numeric(n)
for (i in 1:n) {x_jack <- x[-i]
PCKC_jack[i] <- (quantile(x_jack, 0.75, type = 8) - quantile(x_jack, 0.25, type = 8)) / (2*(quantile(x_jack, 0.90, type = 8) - quantile(x_jack, 0.10, type = 8))) - E_PCK_normal}
acceleration <- sum((mean(PCKC_jack) - PCKC_jack)^3) / (6 * sum((mean(PCKC_jack) - PCKC_jack)^2)^(3/2))
cat("Bootstrap sampling distribution of PCKc statistic (bsd_2$t)", "\n")
agostino.test(bsd_2$t, alternative = "two.sided")
sf.test(bsd_2$t)
ad.test(bsd_2$t)
dagos <- agostino.test(bsd_2$t)
agl <- anscombe.test(bsd_2$t)
ksq_stat <- (dagos$statistic["z"]^2 + agl$statistic["z"]^2)
p_val <- pchisq(ksq_stat, df = 2, lower.tail = FALSE)
cat("D'Agostino-Berlanger-D'Agostino Normality test: bsd_2$t", "\n")
cat("k^2(bsd_2$t) =", round(ksq_stat, 4),",","p_value =", round(p_val, 4), "\n")
t2 <- mean(bsd_2$t)
bsd_2
cat("Bootstrap estimate of PCKc: t* =", t2, "\n")
cat("Jackknife acceleration for PCKc: a =", round(acceleration, 6), "\n")
plot(bsd_2)
boot.ci(bsd_2, conf = 0.95, type = c("norm", "bca", "perc"))
# Bootstrap probability
one_tailed_boot_p <- min(mean(bsd_2$t < 0), mean(bsd_2$t > 0))
two_tailed_boot_p <- 2* one_tailed_boot_p
cat("One-tailed bootstrap probability value for the null hypothesis of kurtosis =", round(one_tailed_boot_p, 4), "\n")
cat("Two-tailed bootstrap probability value for the null hypothesis of kurtosis =", round(two_tailed_boot_p, 4), "\n")
```

5.3. Applied Example: Result of the Script for R With a Sample of 1000 Data Points.

A random sample of 1000 data points following a raised arcsine distribution was generated using the following script, which ensures stability by setting a seed.

```
# Vector x of records of the continuous variable X = waiting time (in days)
set.seed(42)
mu <- 11
s <- 9
n <- 1000
raised_cosine_pdf <- function(wt, mu, s) {
ifelse(wt >= mu - s & wt <= mu + s, 1 / (2 * s) * (1 + cos(pi * (wt - mu) / s)), 0)}
records <- numeric(n)
i <- 1
while (i <= n) {wt <- runif(1, mu - s, mu + s)
y <- runif(1, 0, 1 / s)
if (y <= raised_cosine_pdf(wt, mu, s)) {
records[i] <- wt
i <- i + 1}}
x <- pmin(pmax(records, 2), 20)
print(x)
```

The result of running the script of subsection 5.2 Sub-section with a random sample of 1000 data points generated from a raised arcsine distribution with a location parameter of 11 and a scale parameter of 9 is shown below, including the histogram with the overlaid density and normal curves of the random sample (Figure 5) and histogram and normal Q-Q plot of the bootstrap sampling distribution of the statistic *PCK* (Figure 6).

Wald-Wolfowitz runs test. Criterion: median
Number of runs: r = 488
n_0 = 500, n_1 = 500, and n = n_0 + n_1 = 1000
Two-tailed exact probability value: p = 0.429
Standardized number of runs: z_r = -0.8226
Two-tailed asymptotic probability value: p-value = 0.4107
Grubbs test for one outlier
data: x
G = 2.7826, U = 0.9922, p-value = 1
D'Agostino skewness test: skew = 0.0273, z = 0.3545, p-value = 0.7229
Anscombe-Glynn kurtosis test: kurt = 2.5598, z = -3.5848, p-value = 0.0003373
Anderson-Darling normality test: AD = 0.53, p-value = 0.1755
Lilliefors (Kolmogorov-Smirnov) normality test: D = 0.0212, p-value = 0.336
Shapiro-Francia normality test: W = 0.9965, p-value = 0.0257
D'Agostino-Berlanger-D'Agostino Normality test: k^2(x) = 12.9767, p_value = 0.0015

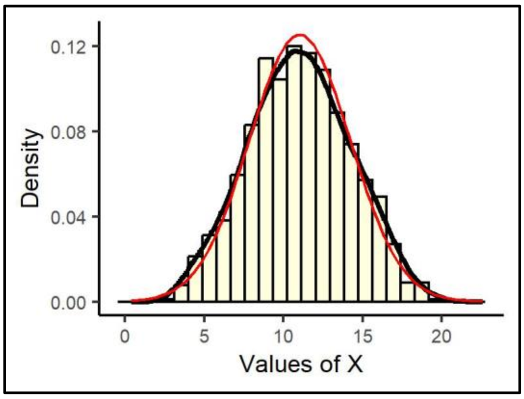


Figure-5. Histogram with overlaid density (black line) and normal (red line) curves of the 1000-data random sample generated from a raised cosine distribution and used as an example of script application.

Kelley's Percentile Coefficient of Kurtosis: PCK = 0.2652
Asymptotic standard error of PCK and PCKc: ase(PCK) = ase(PCKc) = 0.0088
Asymptotic confidence interval 95 % for *PCK*: 95% *CI* (0.2480, 0.2824)
PCK value corresponding to a normal distribution (mesokurtosis) = 0.2632.
Bootstrap sampling distribution of *PCK* statistic (bsd_1$t)
D'Agostino skewness test: skew = -0.0090, z = -0.1165, p-value = 0.9073
Shapiro-Francia normality test: W = 0.9994, p-value = 0.9775
Anderson-Darling normality test: A = 0.2075, p-value = 0.8668
D'Agostino-Berlanger-D'Agostino Normality test: k^2 = 1.2753, p_value = 0.5285

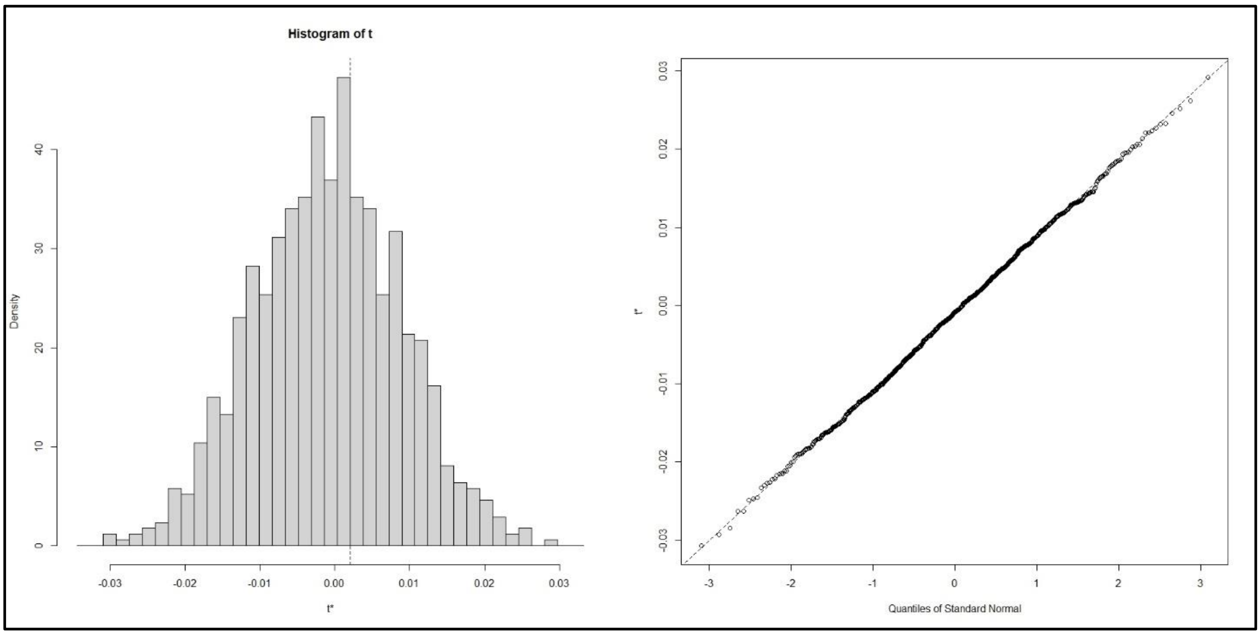


Figure-6. Histogram and normal Q-Q plot of the bootstrap sampling distribution derived from the 1000-data random sample generated from a raised cosine distribution and used as an example of script application.

Bootstrap Statistics:

t*	original	bias	std. error	acceleration
0.2622	0.2652	-0.0030	0.0097	0.0017

Bootstrap confidence interval calculations

Level	Normal	Percentile	BCa
95%	(0.2492, 0.2873)	(0.2438, 0.2815)	(0.2492, 0.2869)

Percentile coefficient of kurtosis centered at 0 = 0.0020
Asymptotic confidence interval 95 % for PCKc: 95% CI (-0.0152, 0.0193)

Bootstrap sampling distribution of PCKc statistic (bsd_2$t)
D'Agostino skewness test: skew = -0.0090, z = -0.1165, p-value = 0.9073
Shapiro-Francia normality test: W = 0.9994, p-value = 0.9775
Anderson-Darling normality test: A = 0.2075, p-value = 0.8668
D'Agostino-Berlanger-D'Agostino Normality test: k^2 = 1.2753, p_value = 0.5285

Bootstrap Statistics:

t*	original	bias	std. error	acceleration
-0.0010	0.0020	-0.0030	0.0097	0.0017

Bootstrap confidence interval calculations

Level	Normal	Percentile	BCa
95%	(-0.0139, 0.0241)	(-0.0194, 0.0184)	(-0.0139, 0.0237)

One-tailed bootstrap probability value for the null hypothesis of kurtosis = 0.461
Two-tailed bootstrap probability value for the null hypothesis of kurtosis = 0.922

At the 5% significance level, the null hypothesis of randomness in the sample data sequence is maintained by the Wald-Wolfowitz test (*runs* = 488, *z* = -0.823, *p*-value = 0.411). The sample comes from a symmetric distribution according to D'Agostino's test ($\sqrt{b_1}$ = 0.027, *z* = 0.355, *p*-value = 0.723), but it is platykurtic as per the Anscombe-Glynn test (b_2 = 2.560, z = -3.585, *p*-value < 0.001), hence lacks outliers by Grubbs' test (*G* = 2.783, *u* = 0.992, *p*-value = 1). It deviates from normality according to the Shapiro-Francia (w = 0.997, p-value = 0.026) and D'Agostino-Berlanger-D'Agostino (k^2 = 12.977, *p*-value = 0.002) tests. However, the Anderson-Darling (*AD* = 0.53, *p*-value = 0.176) and Kolmogorov-Smirnov-Lilliefors (*d* = 0.021, *p*-value = 0.336) tests do not have the specificity to reject the null hypothesis of normality for a distribution as close to the normal distribution as the raised cosine distribution. The similarity between these two distributions can be seen in Figure 5, in which the density curve presents a slightly more flattened profile than the corresponding normal curve, with wider shoulders and shortened tails.

The *PCK* value is slightly above the expected value for the statistic in case the variable follows a normal distribution (*PCK* = 0.265 > 0.263), which, if significant, would indicate platykurtosis. Due to symmetry, absence of outliers, finite moments, proximity to the normal distribution, and large sample size, the asymptotic confidence interval is adequate and includes the expected value of 0.263 for mesokurtosis: 95% *CI* (0.248, 0.2824). Additionally, the bootstrap confidence intervals from all three methods include the expected value. As the bootstrap sampling distribution of the *CPK* statistic follows a normal distribution, the Gaussian method is the most appropriate, with the other two methods not being invalid. Normality is supported by the Shapiro-Francia and D'Agostino-Berlanger-D'Agostino tests at a significance level of 5%. Furthermore, the bell-shaped profile of a normal curve can be observed in the histogram, and the 45° alignment of the point cloud in the normal QQ plot. Congruently, the asymptotic and bootstrap standard errors (*ase* = 0.0088 versus *bse* = 0.0097) are very similar, with a difference of 0.0009.

PCKc offers the same interpretation of mesokurtosis, as it merely involves a change of origin that centers the measure at 0, making the interpretation of the confidence intervals more evident. Thus, the asymptotic interval and the three bootstrap confidence intervals include 0. Furthermore, the symmetry and normality tests, the bootstrap estimation bias, the bootstrap standard error, the jackknife acceleration, and the plots (histogram and QQ normal plot) of the bootstrap sampling distribution of the *PCKc* statistic yield exactly the same results.

Regarding on the interpretation of mesokurtosis, the bootstrap probability maintains the null hypothesis (H_0: *PCKc* = 0) in a two-tailed test (bootstrap p-value = 0.922 > α = 0.05).

The kurtosis coefficient based on the standardized fourth moment (Pearson, K., 1895; Pearson, E., 1931) falls significantly below 3: b_2 = 2.560, 95% asymptotic *CI* (2.258, 2.861), *ase* = 0.154; 95% bootstrap normal *CI* (2.402, 2.712), *bias* = 0.003, *bse* = 0.079, jackknife *a* = 0.004. The kurtosis coefficient based on the standardized fourth cumulant (Fisher, 1930) also falls significantly below 0: g_2 = -0.436, 95% asymptotic *CI* (-0.739, -0.134), *ase* = 0.155; 95% bootstrap normal *CI* (-0.595, -0.283), *bias* = 0.003, *bse* = 0.080, jackknife *a* = 0.004. Similarly, Hogg's (1974) robust adaptive measure of kurtosis takes a value significantly lower than the expected 1.84 for mesokurtosis: *ARK* = 1.744, 95% bootstrap normal *CI* (1.708, 1.779), *bias* = 0.0005, *bse* = 0.018, jackknife *a* = 0.002. Therefore, all three measures indicate platykurtosis, confirmed by corresponding bootstrap probabilities ($p_{2\text{-tailed}}$ = 0 for the three statistics: b_2, g_2, and *ARK*). It should be noted that, in all three cases, the bootstrap sampling distribution of the statistic is symmetric by the D'Agostino test and conforms to normality by the Shapiro-Francia test. The Appendix includes the script for the point and interval estimation of Pearson's b_2, Fisher's g_2, and Hogg's *ARK* coefficients with a significance level of 5%. The results are rounded to four decimal places, except for the acceleration which is rounded to six decimal places. The corresponding instructions are highlighted in blue to modify them if necessary.

6. Discussion

Regarding the first objective, the sampling distribution of Kelley's percentile measure of kurtosis does converge to normality with normal distributions and symmetric distributions of finite moments, such as the uniform and raised cosine distributions (platykurtic) and the Laplace distribution (leptokurtic), as evidenced by the sampling distribution obtained by the bootstrap method (Nakayama and Tuffin, 2024). Based on the central limit theorem (Waudby-Smith *et al.*, 2021), with a large sample (preferably 1000 or more), even if the distribution is asymmetric, as long as it has finite moments, at least up to the fourth moment, this convergence will occur. This would be the case for distributions like the exponential, Pareto, lognormal, and logLaplace distributions, among others (Montgomery and Runger, 2020). It should be noted that in a small or medium-sized sample with an unknown distribution, this convergence is not guaranteed (Zhang *et al.*, 2022).

With respect to the second objective, a script was developed for the R program to estimate Kelley's measure of kurtosis both pointwise and by interval. This includes the original range from 0 to 0.5 (with a mesokurtosis value of 0.263) and the version centered at 0 (mesokurtosis) with a range from -0.263 to 0.237, which is more straightforward to interpret. Since the correction is a change of origin, calculations related to the asymptotic standard error, the statistics and plots of the bootstrap sampling distribution, as well as the bootstrap probability for the null hypothesis of mesokurtosis, are redundant.

The script allows for checking the randomness, symmetry, and kurtosis of the original sample, plotting the distribution using a histogram with overlaid density and normal curves, and assessing the skewness and normality of the bootstrap sampling distribution to decide which type of confidence interval to use. If the original sample is normal, the asymptotic confidence interval is a good choice. Without normality, the alternative is the bootstrap confidence interval. If the bootstrap sampling distribution conforms to normality, the Gaussian method is recommended. On the other hand, if this bootstrap distribution is not normal, but the bias and jackknife acceleration are small and there is symmetry, the percentile method is a good option. Otherwise, the BCa percentile method should be chosen. When more than one method applies, the one that offers the confidence interval with the smallest amplitude can be selected.

The example of script application reveals that the percentile measure of kurtosis, while robust to parametric assumptions by relying on quantiles computed by R's rule 8, is not as specific as classical measures based on the standardized fourth central moment (Pearson, 1895) and standardized fourth cumulant (Fisher, 1930), or even Hogg's robust adaptive measure based on partial means (corresponding to the bottom and top 20% and 50% of ascending ordered scores), especially with distributions close to normal. Therefore, a script for calculating these measures is included in the Appendix. It is noteworthy that the raised cosine distribution was proposed as a model for the distribution of errors by Lagrange (1775/1869), predating Laplace's formulation of the normal distribution, presented to the Royal Academy of Sciences in Paris in 1778 and published two years later in 1781 (Weatherford, 2022).

Building upon the example presented, there is scope for further research into the sensitivity, specificity, and efficacy of various kurtosis measures (Eberl and Klar, 2024; Fisher, 1930; Hogg, 1974; Pearson, 1895) in detecting deviations from mesokurtosis. This includes distributions close to the normal distribution (e.g., semicircular, symmetric triangular, raised cosine, Student's t with at least five degrees of freedom, Fisher's Z, and logistic), as well as those farther from normality characterized by pronounced kurtosis (e.g., uniform, Laplace, and Cauchy) or a combination of kurtosis and skewness (e.g., exponential, Pareto, lognormal, loglogistic, and logLaplace).

Compared to Pearson's and Fisher's measures of kurtosis, Hogg's measure is more suitable for distributions exhibiting positive skewness and leptokurtosis, such as Laplace, Erlang, exponential, power, and Weibull distributions (Peterson, 2022). Bono *et al.* (2020) explored the bias, precision, and accuracy of Hogg's adaptive robust measure of kurtosis (utilizing means at 5% versus 20% in the numerator) and Fisher's g_2 statistic in relation to three leptokurtic distributions with positive skewness (exponential, lognormal, and gamma) across five sample sizes (50, 100, 400, 1000, and 5000). Their findings indicated that Hogg's measure with means at 20% in the numerator performed best. It should be noted that no simulation studies compare the Kelley's measure with other kurtosis measures.

7. Conclusions

Kelley's percentile measure of kurtosis exhibits a sampling distribution that converges to the normal distribution, making Kelley's provided standard error valuable for computing asymptotic Wald-type confidence intervals. However, it appears less sensitive to minor deviations from mesokurtosis compared to the classical Pearson and Fisher measures, as well as Hogg's robust adaptive measure. The use of the script and the Appendix developed for the point and interval estimation of these measures is recommended, which can be useful both practically and didactically. It is suggested that studies be carried out to assess the sensitivity, specificity, and precision of this kurtosis measure in comparison with those of Pearson, Fisher, Hogg, and Eberl and Klar.

References

Ahsanullah, M., Shakil, M. and Kibria, B. G. (2019). On a generalized raised cosine distribution: some properties, characterizations and applications. *Moroccan Journal of Pure and Applied Analysis,* 5(1): 63-85. Available: https://doi.org/10.2478/mjpaa-2019-0006

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics,* 32(2): 193-212. Available: http://dx.doi.org/10.1214/aoms/1177729437

Anscombe, F. J. and Glynn, W. J. (1983). Distribution of kurtosis statistics b2 for normal samples. *Biometrika,* 70(1): 227-34. Available: https://doi.org/10.1093/BIOMET/70.1.227

Baumgarten, C. and Patel, T., 2022. "Automatic random variate generation in Python. In M. Agarwal, C. Calloway, D. Niederhut and D. Shupe (Eds.)." In *Proceedings of the 21st Python in Science Conference. SciPy Organizers.* pp. 46-51.

Bono, R., Arnau, J., Alarcón, R. and Blanca, M. J. (2020). Bias, precision and accuracy of skewness and kurtosis estimators for frequently used continuous distributions. *Symmetry*, 12(1): 19. Available: https://doi.org/10.3390/sym12010019

Bruce, P., Bruce, A. and Gedeck, P. (2020). *Practical statistics for data scientists: 50+ essential concepts using r and python.* O'Reilly Media:

Caeiro, F. and Mateus, A. (2024). randtests: Testing randomness in R. Available: https://doi.org/10.32614/CRAN.package.randtests

Canty, A., Ripley, B. and Brazzale, A. R. (2024). Package 'boot. Available: https://cran.r-project.org/web/packages/boot/boot.pdf

Chattamvelli, R. and Shanmugam, R. (2021). *Cosine distribution. In continuous distributions in engineering and the applied sciences – part i. Synthesis lectures on mathematics and statistics.* Springer. 81–91. https://doi.org/10.1007/978-3-031-02430-6_7

Chattamvelli, R. and Shanmugam, R. (2023). *Descriptive statistics for scientists and engineers. Applications in R.* Springer. https://doi.org/10.1007/978-3-031-32330-0

Chowdhury, S. (2023). *Monte carlo methods utilizing mathematica®: Applications in inverse transform and acceptance-rejection sampling.* Springer Nature. https://doi.org/10.1007/978-3-031-23294-7

D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g1. *Biometrika,* 57(3): 679–81. Available: https://doi.org/10.1093/biomet/57.3.679

D'Agostino, R. B. and Pearson, E. S. (1973). Test for departure from normality. Empirical result for the distribution of b2 and √b1. *Biometrika,* 60(3): 613-22. Available: https://doi.org/10.1093/biomet/60.3.613

D'Agostino, R. B., Berlanger, A. and D'Agostino, R. B. J. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistics,* 44(4): 316-21. Available: https://doi.org/10.2307/2684359

Demir, S. (2022). Comparison of normality tests in terms of sample sizes under different skewness and kurtosis coefficients. *International Journal of Assessment Tools in Education,* 9(2): 397-409. Available: https://doi.org/10.21449/ijate.1101295

Denis, D. J. (2020). *Univariate, bivariate and multivariate statistics using r: Quantitative tools for data analysis and data science.* John Wiley and Sons. https://doi.org/10.1002/9781119549963

Eberl, A. and Klar, B. (2024). Centre-free kurtosis orderings for asymmetric distributions. *Statistical Papers,* 65(1): 415-33. Available: https://doi.org/10.1007/s00362-023-01403-6

Efron, B. (2022). *Exponential families in theory and practice.* Cambridge University Press. https://doi.org/10.1017/9781108773157

Efron, B. and Narasimhan, B. (2020). The automatic construction of bootstrap confidence intervals. *Journal of Computational and Graphical Statistics,* 29(3): 608-19. Available: https://doi.org/10.1080/10618600.2020.1714633

Efron, B. and Narasimhan, B. (2022). Package 'bcaboot'. Bias corrected bootstrap confidence intervals. Available: https://cran.r-project.org/web/packages/bcaboot/bcaboot.pdf

Epanechnikov, V. A. (1969). Nonparametric estimation of a multidimensional probability density. *Theory of Probability and its Applications,* 14(1): 156-61. Available: https://doi.org/10.1137/1114019

Favero, L. P. and Belfiore, P. (2019). *Data science for business and decision making.* Academic Press: Cambridge, MA.

Fisher, R. A. (1930). The moments of the distribution for normal samples of measures of departure from normality. *Proceedings of the Royal Society of London,* 130(812): 16-28. Available: https://doi.org/10.1098/rspa.1930.0185

Freedman, D. and Diaconis, P. (1981). On the histogram as a density estimator: L2 theory. *Probability Theory and Related Fields,* 57(4): 453-76. Available: https://doi.org/10.1007/BF01025868

Gross, J. and Ligges, U. (2022). Package 'nortest. Available: https://cran.r-project.org/web/packages/nortest/nortest.pdf

Grubbs, F. (1969). Procedures for detecting outlying observations in samples. *Technometrics,* 11(1): 1-21. Available: https://doi.org/10.1080/00401706.1969.10490657

Ho, A. D. and Yu, C. C. (2015). Descriptive statistics for modern test score distributions: Skewness, kurtosis, discreteness and ceiling effects. *Educational and Psychological Measurement,* 75(3): 365-88. Available: https://doi.org/10.1177/0013164414548576

Hogg, R. V. (1974). Adaptive robust procedures: a partial review and some suggestions for future applications and theory. *Journal of the American Statistical Association,* 69(348): 909-23. Available: https://doi.org/10.2307/2286160

IBM Corporation (2023). IBM SPSS statistics for windows. Version 29.0.2.0. Ibm corporation. Available: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-29020

Ihaka, R. (2009). The R project: a brief history and thoughts about the future. *Ross Ihaka's Home Page*: Available: https://www.stat.auckland.ac.nz/~ihaka/downloads/Interface98.pdf

Jiang, Z., Ling, N., Lu, Z., Tjostheim, D. and Zhang, Q. (2020). On bandwidth choice for spatial data density estimation. *Journal of the Royal Statistical Society Series B: Statistical Methodology,* 82(3): 817-40. Available: https://doi.org/10.1111/rssb.12367

Kelley, T. L. (1923). *Statistical method.* The Macmillan Company.

Kelley, T. L. (1947). *Fundamentals of statistics.* Harvard University Press.

Kolmogorov, A. N. (1933). Sulla determinizione empirica di una legge di distribuzione. *Giornale Dell Istituto Italiano Degli Attuari,* 4: 83-91.

Komsta, L. (2022). outliers: Tests for outliersv. Available: https://doi.org/10.32614/CRAN.package.outliers

Komsta, L. and Novomestky, F. (2022). moments: Moments, cumulants, skewness, kurtosis and related tests. Available: https://doi.org/10.32614/CRAN.package.moments

Kyoungjae, L., Lizhen, L. and Kisung, Y. (2022). Statistical hypothesis testing tool. Boxvar1.Chisq {sht}. One-sample chi-square test for variance. In r documentation. Available: https://search.r-project.org/CRAN/refmans/SHT/html/var1.chisq.html

Lagrange, J. L. (1775/1869). Recherches sur les suites récurrentes dont les termes varient de plusieurs manières différentes, ou sur l'integration des équations linéaires aux différences finies et partielles, et sur l'usage de ces équations dans la théorie des hasards. In J. A. Serret (Ed.), Oeuvres Complètes (vol. 4, pp. 151−251). Gauthier–Villars. Original publication in Nouveau. *Memoires de l'Academie Royale des Sciences et Belles-Lettres de Berlin,* 6: 182-272.

Lakens, D. (2022). Sample size justification. *Collabra: Psychology,* 8(1): 33267. Available: https://doi.org/10.1525/collabra.33267

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknow. *Journal of the American Association,* 62(318): 399-402. Available: https://doi.org/10.2307/2283970

Linden, A. (2023). *CENTILE2: Stata module to enhance centile command and provide additional definitions for computing sample quantiles. Statistical software components. S459262.* Boston College Department of Economics. https://www.gnu.org/licenses/gpl-3.0.txt

McIntyre, D. and Chow, C. K. (2020). Waiting time as an indicator for health services under strain: a narrative review. INQUIRY. *The Journal of Health Care Organization, Provision and Financing,* 57: 0046958020910305. Available: https://doi.org/10.1177/0046958020910305

Montgomery, D. C. and Runger, G. C. (2020). *Applied statistics and probability for engineers.* John Wiley and Sons.

Moral, J. (2023). Shape measures for the distribution of a qualitative variable. *Open Journal of Statistics,* 13(4): 619-34. Available: https://doi.org/10.4236/ojs.2023.134030

Moral, J. (2024). Determination of the number and width of class intervals using R. *Annals of Environmental Science and Toxicology,* 18(1): 22-42. Available: https://dx.doi.org/10.17352/aest.000077

Nakayama, M. K. and Tuffin, B. (2024). Sufficient conditions for central limit theorems and confidence intervals for randomized quasi-Monte Carlo methods. *ACM Transactions on Modeling and Computer Simulation*, 34(3): 13. Available: https://doi.org/10.1145/3643847

Pearson, K. (1895). Contributions to the mathematical theory of evolution. Ii. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London A*, 186: 343-414. Available: https://doi.org/10.1098/rsta.1895.0010

Pearson, K. (1905). Das fehlergesetz und seine verallgemeinerungen durch fechner und pearson. A rejoinder [The law of error and its generalizations by Fechner and Pearson. A rejoinder]. *Biometrika*, 4(1-2): 169-212. Available: https://doi.org/10.1093/biomet/4.1-2.169

Peterson, L. E. (2022). Skewness and kurtosis of 18 simulated probability distributions with varying sample size and scale. *Social Science Research Network (SSRN)*: 4277814. Available: http://dx.doi.org/10.2139/ssrn.4277814

R Core Team (2020). R: A language and environment for statistical computing. R foundation for statistical computing. Available: https://www.R-project.org/

R Core Team and contributors worldwide (2024). Quantile {stats}. R documentation. Sample quantiles. Available: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/quantile.html

R Studio Team (2020). Rstudio: Integrated development for r. Rstudio. Public benefit corporation. Available: http://www.rstudio.com/

Ramachandran, K. M. and Tsokos, C. P. (2020). *Mathematical statistics with applications in R.* Academic Press.

Royston, J. P. (1993). A toolkit for testing for non-normality in complete and censured samples. *Journal of the Royal Statistical Society, Series D (The Statistician):*, 42(1): 37-43. Available: https://doi.org/10.2307/2348109

Shapiro, S. S. and Francia, R. S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67(337): 215–16. Available: https://doi.org/10.1080/01621459.1972.10481232

Sharipova, M. (2024). Asymmetry and kurtosis coefficients. *Theoretical Aspects in The Formation of Pedagogical Sciences*, 3(1): 216-25.

Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society, Series B. Statistical Methodology*, 53(3): 683–90. Available: https://doi.org/10.1111/j.2517-6161.1991.tb01857.x

Sholl, D. S. and Steckel, J. A. (2022). *Density functional theory: a practical introduction.* John Wiley and Sons.

Smirnov, N. V. (1948). Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19(2): 279-81. Available: https://doi.org/10.1214/aoms/1177730256

Snedecor, G. W. and Cochran, W. G. (1989). *Statistical methods.* 8th ednIowa State University Press.

Sukhoplyuev, D. I. and Nazarov, A. N. (2024). *Methods of descriptive statistics in telemetry tasks. In ieee (ed.), proceedings of the 2024 systems of signals generating and processing in the field of on-board communications (article 10496798).* Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/IEEECONF60226.2024.10496798

Wald, A. (1939). Contributions to the theory of statistical estimation and testing hypotheses. *Annals of Mathematical Statistics*, 10(4): 299-326. Available: https://doi.org/10.1214/aoms/1177732144

Wald, A. and Wolfowitz, J. (1943). An exact test for randomness in the non-parametric case based on serial correlation. *The Annals of Mathematical Statistics*, 14(4): 378–88. Available: https://doi.org/10.1214/aoms/1177731358

Waudby-Smith, I., Arbour, D., Sinha, R., Kennedy, E. H. and Ramdas, A., 2021. "Time-uniform central limit theory and asymptotic." In *Confidence Sequences. Arxiv, Article 2103.06476.*

Weatherford, R. (2022). *Philosophical foundations of probability theory.* Routledge. https://doi.org/10.4324/9781003306993

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D. and van den Brand, T. (2024). Ggplot2: Create elegant data visualisations using the grammar of graphics. Available: https://doi.org/10.32614/CRAN.package.ggplot2

Zepeda-Tello, R., Schomaker, M., Maringe, C., Smith, M. J., Belot, A., Rachet, B., Schnitzer, M. E. and Luque-Fernandez, M. A. (2022). *The delta-method and influence function in medical statistics: a reproducible tutorial.* Arxiv, Article 2206.15310. https://doi.org/10.48550/arXiv.2206.15310

Zhang, X., Astivia, O. L. O., Kroc, E. and Zumbo, B. D. (2022). How to think clearly about the central limit theorem. *Psychological Methods*, 28(6): 1427-45. Available: https://doi.org/10.1037/met0000448

Appendix – Calculation of Pearson's b_2, Fisher's g_2, and Hogg's ARK

```
# Vector of scores. Put numerical scores or records separated by commas.
x <- c(13.551419, 15.258590, …, 5.912899)
# Load library
library(moments)
library(nortest)
library(boot)
# Point estimate of Pearson's b2
n <- length(x)
cat("Sample size: n =",n,"\n")
```

```
m4 <- sum((x - mean(x))^4) / length(x)
sd <- sqrt(sum((x - mean(x))^2) / length(x))
b2 <- m4 / sd^4
cat("Kurtosis measure based on the standardized fourth central moment: b_2 =", round(b2, 4), "\n")
# Asymptotic confidence interval for Pearson's b_2
e_b2 <- 3 * (n - 1) / (n + 1)
ase_b2 <- sqrt((24 * n * (n-2) * (n-3))/((n+1)^2 * (n+3) * (n+5)))
cat("Mathematical expectation of b2: e(b2) =", round(e_b2, 4), "\n")
cat("Asymptotic standard error of b2: ase(b2) =", round(ase_b2, 4), "\n")
LL <- b2 - qnorm(0.975)*ase_b2
UL <- b2 + qnorm(0.975)*ase_b2
cat("The 95% asymptotic confidence interval for b2: (", round(LL, 4), ", ", round(UL, 4), ")\n")
# Bootstrap confidence interval for Pearson's b_2
set.seed(123)
bsd<- boot(data = x, function(x, i) {m4 <- sum((x[i] - mean(x[i]))^4) / length(x[i])
sd <- sqrt(sum((x[i] - mean(x[i]))^2) / length(x[i]))
b2 <- m4 / (sd^4)}, R = 1000)
jack_b2 <- numeric(n)
for (i in 1:n) {jack_x <- x[-i]
m4 <- sum((jack_x - mean(jack_x))^4) / length(jack_x)
sd <- sqrt(sum((jack_x - mean(jack_x))^2) / length(jack_x))
jack_b2[i] <- m4 / (sd^4)}
acceleration <- sum((mean(jack_b2) - jack_b2)^3) / (6 * sum((mean(jack_b2) - jack_b2)^2)^(3/2))
agostino.test(bsd$t, alternative = "two.sided")
sf.test(bsd$t)
bsd
cat("Jackknife acceleration: a =", round(acceleration, 6), "\n")
plot(bsd)
boot.ci(bsd, conf = 0.95, type = c("norm", "bca", "perc"))
# Bootstrap probability for mesokurtosis using Pearson's b_2 statistic
two_tailed_boot_p <- 2 * min(mean(bsd$t < 3 * (n-1) / (n+1)), mean(bsd$t > 3 * (n-1) / (n+1)))
cat("Two-tailed bootstrap probability value for null hypothesis of β2 = 3: p_value =", round(two_tailed_boot_p, 4), "\n")
# Point estimate of Fisher's g_2
g2 <- (n*(n+1)) / ((n-1)*(n-2)*(n-3)) * (sum((x - mean(x))^4) / sd(x)^4) - 3* (n-1)^2 / ((n-2)*(n-3))
cat("Excess kurtosis based on standardized quarter cumulant: g_2 =", round(g2, 4), "\n")
# Asymptotic confidence interval for Fisher's g_2
se_g2 <- sqrt((24 * n * (n-1)^2)/((n-3) * (n-2) * (n+3) * (n+5)))
cat("Asymptotic standard error of g_2: se(g2) =", round(se_g2, 4), "\n")
LL <- g2 - qnorm(0.975)*se_g2
UL <- g2 + qnorm(0.975)*se_g2
cat("The 95% asymptotic confidence interval for g_2: (", round(LL, 4), ", ", round(UL, 4), ")\n")
# Bootstrap confidence interval for Fisher's g_2
set.seed(123)
bsd <- boot(data = x, function(x, i) {n <- length(x[i])
g2 <- (n*(n+1)) / ((n-1)*(n-2)*(n-3)) * (sum((x[i] - mean(x[i]))^4) / sd(x[i])^4) - 3*(n-1)^2/((n-2)*(n-3))}, R = 1000)
jack_g2 <- numeric(n)
for (i in 1:n) {jack_x <- x[-i]
n <- length(jack_x)
jack_g2[i] <- (n*(n+1)) / ((n-1)*(n-2)*(n-3)) * (sum((jack_x - mean(jack_x))^4) / sd(jack_x)^4) - 3*(n-1)^2/((n-2)*(n-3))}
acceleration <- sum((mean(jack_g2) - jack_g2)^3) / (6 * sum((mean(jack_g2) - jack_g2)^2)^(3/2))
agostino.test(bsd$t, alternative = "two.sided")
sf.test(bsd$t)
bsd
cat("Jackknife acceleration: a =", round(acceleration, 6), "\n")
plot(bsd)
boot.ci(bsd, conf = 0.95, type = c("norm", "bca", "perc"))
# Bootstrap probability for mesokurtosis using Fisher's g_2 statistic
two_tailed_boot_p <- 2 * min(mean(bsd$t < 0), mean(bsd$t > 0))
cat("Two-tailed bootstrap probability value for null hypothesis of γ = 0: p_value =", round(two_tailed_boot_p, 4), "\n")
# Point estimate of Hogg's Adaptive Robust Kurtosis (ARK)
xo <- sort(x)
p_20 <- 0.2 * n - floor(0.2 * n)
```

```
l_n_20 <- floor(0.2 * n)
u_n_20 <- n - l_n_20 + 1
lm_20 <- (sum(xo[1:l_n_20]) + p_20 * xo[l_n_20 + 1]) / (n * 0.2)
um_20 <- (sum(xo[u_n_20:n]) + p_20 * xo[u_n_20 - 1]) / (n * 0.2)
p_50 <- 0.5 * n - floor(0.5 * n)
l_n_50 <- floor(0.5 * n)
u_n_50 <- n - l_n_50 + 1
lm_50 <- (sum(xo[1:l_n_50]) + p_50 * xo[l_n_50 + 1]) / (n * 0.5)
um_50 <- (sum(xo[u_n_50:n]) + p_50 * xo[u_n_50 - 1]) / (n * 0.5)
cat("Lower mean (the bottom 20% of the ascendingly sorted data) =", round(lm_20, 4), "\n")
cat("Upper mean (the top 20% of the ascendingly sorted data) =", round(um_20, 4), "\n")
cat("Lower mean (the bottom 50% of the ascendingly sorted data) =", round(lm_50, 4), "\n")
cat("Upper mean (the top 50% of the ascendingly sorted data) =", round(um_50, 4), "\n")
ARK <- (um_20 - lm_20) / (um_50 - lm_50)
cat("Hogg’s Adaptive Robust Kurtosis =", round(ARK, 4), "\n")
# Bootstrap confidence interval for ARK
set.seed(123)
bsd <- boot(data = x, statistic = function(x, i) {n <- length(x[i])
xo <- sort(x[i])
p_20 <- 0.2 * n - floor(0.2 * n)
l_n_20 <- floor(0.2 * n)
u_n_20 <- n - l_n_20 + 1
lm_20 <- (sum(xo[1:l_n_20]) + p_20 * xo[l_n_20 + 1]) / (n * 0.2)
um_20 <- (sum(xo[u_n_20:n]) + p_20 * xo[u_n_20 - 1]) / (n * 0.2)
p_50 <- 0.5 * n - floor(0.5 * n)
l_n_50 <- floor(0.5 * n)
u_n_50 <- n - l_n_50 + 1
lm_50 <- (sum(xo[1:l_n_50]) + p_50 * xo[l_n_50 + 1]) / (n * 0.5)
um_50 <- (sum(xo[u_n_50:n]) + p_50 * xo[u_n_50 - 1]) / (n * 0.5)
(um_20 - lm_20) / (um_50 - lm_50)}, R = 1000)
ARK_jack <- numeric(n)
for (i in 1:n) {x_jack <- x[-i]
xo <- sort(x_jack)
n <- length(x_jack)
p_20 <- 0.2 * n - floor(0.2 * n)
l_n_20 <- floor(0.2 * n)
u_n_20 <- n - l_n_20 + 1
lm_20 <- (sum(xo[1:l_n_20]) + p_20 * xo[l_n_20 + 1]) / (n * 0.2)
um_20 <- (sum(xo[u_n_20:n]) + p_20 * xo[u_n_20 - 1]) / (n * 0.2)
p_50 <- 0.5 * n - floor(0.5 * n)
l_n_50 <- floor(0.5 * n)
u_n_50 <- n - l_n_50 + 1
lm_50 <- (sum(xo[1:l_n_50]) + p_50 * xo[l_n_50 + 1]) / (n * 0.5)
um_50 <- (sum(xo[u_n_50:n]) + p_50 * xo[u_n_50 - 1]) / (n * 0.5)
ARK_jack[i] <- (um_20 - lm_20) / (um_50 - lm_50)}
accel <- sum((mean(ARK_jack) - ARK_jack)^3) / (6 * sum((mean(ARK_jack) - ARK_jack)^2)^(3/2))
agostino.test(bsd$t, alternative = "two.sided")
sf.test(bsd$t)
bsd
cat("Jackknife acceleration: a =", round(accel, 6), "\n")
plot(bsd)
# Bootstrap probability for mesokurtosis using ARK
boot.ci(bsd, conf = 0.95, type = c("norm", "bca", "perc"))
p_boot <- min(mean(bsd$t < 1.84), mean(bsd$t > 1.84))
cat("Two-tailed bootstrap probability value for null hypothesis of mesokurtosis (ARK = 1.84): p =", round(p_boot, 4), "\n").
```