Journal of Statistics & Applied Sciences, Issue -10
ISSN 2718-0999

Research Article

Received: date:04.07.2024
Accepted: date:23.09.2024
Published: date:31.12.2024

Kelley’s coefficients of skewness using R

Jose Moral de la Rubia[1*]

[1]School of Psychology, Universidad Autónoma de Nuevo León, Monterrey Mexico; jose.morald@uanl.edu.mx
Orcid: 0000-0003-1856-1458[1]
*Correspondence: jose.morald@uanl.edu.mx

Abstract: Kelley developed a robust measure of asymmetry based on quantiles. His proposal was an absolute index which, when divided by the median, results in its relative expression. If the additive complement is standardized with the semi-percentile range, the percentile coefficient of skewness (PCS) is obtained. Additionally, Kelley provided its standard error in case of normal distribution. However, no statistical software currently computes these measures. The aim of this methodological article is to determine their sampling distribution and facilitate their use. Three random samples of 10,000 data points were generated from three symmetric distributions: semicircular (platykurtic), normal (mesokurtic), and logistic (leptokurtic). By bootstrapping, the sampling distribution was obtained for absolute and relative indices, as well as the PCS. The sampling distributions of the absolute index and the PCS conformed to normality, while that of the relative index was leptokurtic with an excessive bootstrap standard error. Furthermore, a script was developed for the R program, adjusted based on these findings, to obtain point and interval estimates of these indices. The script was applied to a random sample as an example. It is concluded that dividing the absolute index by the semi-percentage range is a better standardization option than dividing by the median.

Keywords: skewness; quantiles; bootstrap confidence interval; asymptotic standard error; normality.

1. Introduction

In 1923, the American psychologist Truman Lee Kelley (1884-1961) proposed a robust measure of skewness based on quantiles, which he had initially proposed in 1921 as a measure of dispersion [1]. The 1923 formulation was an absolute index that, when divided by the median, results in its relative expression, which is not bounded between -1 and 1 [2]. Moreover, his original measure of skewness is interpreted in the opposite way compared to the classical coefficients of Pearson [3-4], Bowley [5], Yule [6], and Fisher [7], so it went somewhat unnoticed, except among researchers in the field of psychology [8]. However, if the additive complement of the absolute index is standardized by dividing it by the semi-percentile range, the percentile coefficient of skewness is obtained, which is bounded between -1 and 1. Additionally, Kelley [2] provided the standard error of the absolute index for a normal distribution, which can be adapted to its two standardized expressions, whether divided by the absolute value of the median or by the semi-percentile range.

It has not been encountered so far, no statistical program computes this measure. The objective of this methodological article is to determine the sampling distribution of the absolute index and the two relative modalities proposed, as well as to facilitate the use of these indices by means of a script developed for the R program, which has been developed by the mathematical community and is freely available [9]. This script allows for the point and interval estimation, providing the information required to choose between the asymptotic or bootstrap confidence interval and, in the latter case, to decide among three estimation methods: normal, percentile, and bias-corrected and accelerated percentile. For this purpose, the assumptions of randomness and normality of the sample are tested. The sample is represented by a histogram with overlaid density and normal curves. In addition, the symmetry, skewness, and normality of the bootstrap sampling distribution are checked.

Citation : J. M. de la Rubia, "Kelley’s coefficients of skewness using R," *Journal of Statistics and Applied Sciences*, no. 10, pp. 50–75, Dec. 2024, doi:10.52693/jsas.1510230

2. Kelley’s absolute index of skewness

Kelley's absolute index of skewness is the signed distance from the mean percentile range to the median. It is calculated by subtracting from the median (minuend) the mean percentile range (subtrahend) and can be denoted as *AIS*. See Equation 1, where $q_p(\mathrm{x})$ represents the p-order sample quantile. It is interpreted oppositely to the traditional measures of skewness [3-7], as well as more contemporary measures by Bickel [9], Altinay [10], Singh, Gewali, and Khatiwada [11], and Eberl and Klar [12]. Negative values indicate that the right side of the distribution is longer than the left side or the presence of a long tail on the right, taking the median as the axis of symmetry. Conversely, positive values reveal that the left side of the distribution is longer than the right or the presence of a long tail on the left.

$$\mathrm{x} = \{x_i\}_{i=1}^{n} = \{x_1, x_2, \dots, x_n\} \subseteq \mathrm{X}$$
$$\widehat{AIS}(\mathrm{x}) = d\left(\frac{q_{0.1}(\mathrm{x}) + q_{0.90}(\mathrm{x})}{2}, q_{0.5}(\mathrm{x})\right) = q_{0.5}(\mathrm{x}) - \frac{q_{0.1}(\mathrm{x}) + q_{0.90}(\mathrm{x})}{2} \tag{1}$$

The calculation of sample quantiles can be done by rules 6, 7, 8, or 9 in the R program [14]. The SPSS program calculates quantiles using rule 6, which is the most used method [15]. This rule expresses the order p of the quantile $x_{(p)}$ as the average or expected value of the i-th order statistic in a sample of size n drawn from a standard continuous uniform distribution $U[0, 1]$. This approach is used as a non-informative prior distribution when estimating a probability in Bayesian inference [16]. Rule 7, which is the default in the R program, expresses the order p of the quantile $x_{(p)}$ as the mode of the i-th order statistic in a sample of size n [17]. Rule 8 is especially recommended when the distribution of the variable is unknown [18-19], as suggested by the simulation study of Hyndman and Fan [20] and the exploratory data analysis of Tukey [21]. This rule expresses the order p of the quantile $x_{(p)}$ as the median of the i-th order statistic in a sample of size n drawn from a standard continuous uniform distribution. For normal distributions, rule 9 [22] is recommended. In this rule, the quantile order is expressed as $p = (i - a) / (n + 1 - a - b)$, where $a = b = 3/8$, n is the sample size, and i denotes the unknown order of the quantile among the n sample data, so $i = 0.375 + p \times (n + 0.25)$. If i is an integer, after sorting the sample in ascending order, the data in that order is the quantile. If i is a number with decimals, the quantile is obtained by linear interpolation (Equation 2).

$$\mathrm{x} = \{x_i\}_{i=1}^{n} = \{x_1, x_2, \dots, x_n\} \subseteq \mathrm{X}$$
$$x_{(1)} \le x_{(2)} \le \cdots \le x_{(i)} \le \cdots \le x_{(n)}$$
$$p = \frac{i - a}{n + 1 - a - b} \Rightarrow \begin{array}{lcc} Rule\ 6 & a = b = 0 & i = p(n+1) \\ Rule\ 7 & a = b = 1 & i = 1 + p(n-1) \\ Rule\ 8 & a = b = 1/3 & i = 1/3 + p(n + 1/3) \\ Rule\ 9 & a = b = 3/8 & i = 0.375 + p(n + 0.25) \end{array} \tag{2}$$
$$q_p(\mathrm{x}) = \begin{cases} x_{(i)} & i = \lfloor i \rfloor = \lceil i \rceil \\ x_{(\lfloor i \rfloor)} + (i - \lfloor i \rfloor)\left(x_{(\lfloor i \rfloor + 1)} - x_{(\lfloor i \rfloor)}\right) & i = \lfloor i \rfloor + (i - \lfloor i \rfloor) \end{cases}$$

In addition, Kelley [1-2] provided a formula for the standard error (*se*) or standard deviation in the sampling distribution of the *AIS*, assuming a normal distribution. See Equation 3, where $PR(\mathrm{x})$ is the sample percentile range, $q_p(\mathrm{x})$ represents the sample quantile of order p, and n denotes the sample size. The constant 0.599143 is a scaling factor derived from theoretical calculations that model the dispersion of the inter-percentile range, divided by the square root of the sample size, in random samples from a standard normal distribution. This factor ensures that the calculated standard error aligns with the asymptotic behavior of percentiles in random samples of varying sizes drawn from a standard normal distribution [1].

$$\mathrm{x} = \{x_i\}_{i=1}^{n} = \{x_1, x_2, \dots, x_n\} \subseteq \mathrm{X} \sim N(0,1)$$
$$se\left[\widehat{AIS}(\mathrm{x})\right] = 0.599143 \times \frac{PR(\mathrm{x})}{\sqrt{n}} = 0.599143 \times \frac{q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x})}{\sqrt{n}} \tag{3}$$

For a random sample drawn from a normal distribution, or a large sample drawn from a variable with finite mean and variance, this formula for the standard error enables the asymptotic estimation of the confidence interval at a confidence level of $(1 - \alpha) \times 100$. See Equation 4, where LL represents the low limit and UL denotes the upper limit of the confidence interval.

$$P(LL \le AIS \le UL) = P(AIS(\mathrm{X}) \in [LL, UL]) = 1 - \alpha$$
$$LL = \widehat{AIS}(\mathrm{x}) - z_{1-\frac{\alpha}{2}} \times se\left[\widehat{AIS}(\mathrm{x})\right] \tag{4}$$

$$UL = \widehat{AIS}(\mathrm{x}) + z_{1-\frac{\alpha}{2}} \times se\left[\widehat{AIS}(\mathrm{x})\right]$$

$$z_{1-\frac{\alpha}{2}} = \Phi^{-1}\left(1-\frac{\alpha}{2}\right); \Phi^{-1}\left(1-\frac{0.05}{2}\right) = \Phi^{-1}(0.975) \approx 1.96$$

In turn, for any type of distribution or a random variable with an unknown distribution, the confidence interval can be calculated using bootstrap with the extraction of 1000 or more samples, applying the bias-corrected and accelerated percentile, percentile, or normal method, unless the bootstrap standard error is too large. The bootstrap standard error should preferably be less than one-fourth of the sample interval. A value greater than half of the sample range indicates a very large error [23]. Additionally, for the method to be accurate, a random sample that is representative of the population and preferably not less than 30 data points is required. If the bootstrap sampling distribution of the skewness statistic follows a normal distribution, the normal method can be used, though the other two methods are not necessarily inadequate. If the distribution is symmetric (tested by the D'Agostino test [24], or another statistical test) with small skewness and acceleration ($|bias| \leq 0.05$ and $|a| \leq 0.025$, respectively), the percentile method can be applied, with the bias-corrected and accelerated method not being inappropriate. Otherwise, the bias-corrected and accelerated percentile method is the best choice [25-26]. When more than one method is applicable, the one that provides the confidence interval with the smallest width can be chosen, as efficiency is a desirable property.

3. Kelley's relative index of skewness

The so-called Kelley's relative index of skewness is obtained by dividing the absolute index by the median. This results in a quotient with the difference between the median and the mean percentile range in the numerator and the median in the denominator (Equation 5). Thus, it is standardized in a manner analogous to relative ranges. However, like these measures of variability, the denominator does not convert the quotient into a ratio with a range from -1 to 1, but instead allows it to vary from -∞ to ∞ [27]. This index can be denoted as *RIS*.

$$\widehat{RIS}(\mathrm{x}) = \frac{q_{0.5}(\mathrm{x}) - \frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x})}{2}}{q_{0.5}(\mathrm{x})} = 1 - \frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x})}{2 \times q_{0.5}(\mathrm{x})} \tag{5}$$

If the sampling distribution of the *RIS* were a normal distribution and the median were considered a constant instead of a random variable, the standard error of the *AIS* would be multiplied by the inverse of the absolute value of this constant, resulting in the standard error of the relative index (Equation 6). Consequently, an asymptotic confidence interval could be defined, as shown in Equation 4.

$$se\left[\widehat{RIS}(\mathrm{x})\right] = se\left[\frac{\widehat{AIS}(\mathrm{x})}{q_{0.5}(\mathrm{x})}\right] = \frac{1}{|q_{0.5}(\mathrm{x})|} se\left[\widehat{AIS}(\mathrm{x})\right] = \frac{0.599143}{\sqrt{n}} \times \frac{PR(\mathrm{x})}{|q_{0.5}(\mathrm{x})|} \tag{6}$$

To maintain the interpretative logic of the skewness coefficients established by Pearson [3-4] and settled by Bowley [5] and Yule [6], the order of the minuend and subtrahend in *AIS* can be reversed (Equation 7). The index value remains the same, with only the sign changing. A positive sign in the additive opposite of *AIS* indicates positive or right-tailed skewness, while a negative sign indicates negative or left-tailed skewness. This reversal can also be applied to the additive opposite of *RIS*, changing only the sign (Equation 8). Additionally, the value of the asymptotic standard error remains unchanged in either index (Equations 9 and 10). These additive opposites can be denoted by *AISr* and *RISr*, where *r* represents the revised or reversed character of the index.

$$1 - AIS(\mathrm{x}) = \frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x})}{2} - q_{0.5}(\mathrm{x}) = \frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x}) - 2 \times q_{0.5}(\mathrm{x})}{2} = AIS_r(\mathrm{x}) \tag{7}$$

$$1 - RIS(\mathrm{x}) = \frac{\frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x})}{2} - q_{0.5}(\mathrm{x})}{q_{0.5}(\mathrm{x})} = \frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x})}{2 \times q_{0.5}(\mathrm{x})} - 1 = RIS_r(\mathrm{x}) \tag{8}$$

$$se[1 - AIS(\mathrm{x})] = se[AIS(\mathrm{x})] = 0.599143 \times \frac{PR(\mathrm{x})}{\sqrt{n}} = se[AIS_r(\mathrm{x})] \tag{9}$$

$$se[1 - RIS(\mathrm{x})] = se[RIS(\mathrm{x})] = 0.599143 \times \frac{PR(\mathrm{x})}{\sqrt{n}} = se[RIS_r(\mathrm{x})] \tag{10}$$

In the latter case, if instead of standardizing using the median, it is divided by the semi-percentile range, as in Altinay's measure of skewness [11], it results in the percentile coefficient of skewness as the relative skewness index. This index is a ratio and is bounded between -1 and 1 (Equation 11).

$$RIS_r(\mathrm{x}) = \frac{\frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x})}{2} - q_{0.5}(\mathrm{x})}{\frac{q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x})}{2}} = \frac{q_{0.1}(\mathrm{x}) + q_{0.9}(\mathrm{x}) - 2 \times q_{0.5}(\mathrm{x})}{q_{0.9}(\mathrm{x}) - q_{0.1}(\mathrm{x})} = PCS(\mathrm{x}) \quad (11)$$

If the sampling distribution of the percentile coefficient of skewness were normal and the percentile range were considered a constant instead of a random variable, the standard error would be simplified to the expression shown in Equation 12, which would allow the definition of an asymptotic confidence interval as shown in Equation 4.

$$se\left(\frac{1 - AIS(\mathrm{x})}{PR(\mathrm{x})}\right) = \frac{1}{PR(\mathrm{x})} se[AIS(\mathrm{x})] = \frac{0.599143 \times PR(\mathrm{x})/\sqrt{n}}{PR(\mathrm{x})/2} = \frac{1.19828}{\sqrt{n}} \quad (12)$$

It should be noted that the percentile coefficient of skewness does not have a clear attribution of authorship. In 1975, Hinkley [28] used this coefficient to assess symmetry after applying the Box-Cox power transformation to sample data drawn from non-normal distributions, marking the first published use of this measure of skewness. In the field of economics [29-30] and other disciplines [31-32], it is attributed to Kelly, often without proper citation, likely causing confusion with the psychologist Truman Lee Kelley [33-34].

The percentile coefficient of skewness (*PCS*) deals with one of the infinite possibilities of Bowley's [5] generalization of the quartile coefficient of skewness given by Groeneveld and Meeden [35]. It is similar to the quartile coefficient of skewness except that it uses the 10th ($q_{0.1}$) and 90th ($q_{0.9}$) percentiles instead of the 1st ($q_{0.25}$) and 3rd ($q_{0.75}$) quartiles in the numerator and uses the percentile range (*PR*) instead of the interquartile range (*IQR*) in the denominator. Thus, *PCS* is a way of standardizing the additive opposite of Kelley's *AIS*.

Using quantiles, *AIS* is a robust measure of skewness and can be applied to any type of distribution. The percentile coefficient of skewness is considered a better choice than the quartile coefficient of skewness by Bowley [5] and Yule [6] for platykurtic or thin-tailed distributions, such as the uniform, triangular, or semicircular distribution [29, 36].

4. Materials and Methods

For the first objective of the study, which was to determine the sampling distribution of the three measures of skewness: *AISr*, *RISr*, and *PCS*, three random samples of 10,000 data points were generated using the inverse transform method [37]. Three continuous, symmetric distributions with finite moments were chosen to facilitate compliance with the central limit theorem [17], each with different kurtosis. The Wigner's standard semicircular distribution $SC(r = 1)$ was selected as the platykurtic distribution. This was generated from a beta distribution by applying the transformation: $Y = 2rX - r = 2X - 1$, where the variable Y follows a semicircular distribution with unit radius ($r = 1$) and the variable X follows a Beta distribution with shape parameters: $\alpha = 2/3$ and $\beta = 3/2$ [38]. The standard normal distribution with location parameter 0 and scale 1, $N(0, 1)$. was taken as the mesokurtic distribution. The chosen leptokurtic distribution was the standard logistic distribution with location parameter 0 and scale 1, $Logist(0, 1)$.

From each of these three source samples of 10,000 data points, 1,000 samples of 10,000 data points were drawn with replacement. In each of these bootstrap samples, Kelley's absolute and relative indices, as well as the percentile coefficient of skewness, were calculated. Thus, three bootstrap sampling distributions were obtained for each of the three measures of skewness.

For each of these nine bootstrap sampling distributions, the presence of outliers was checked using Grubbs' test [39], symmetry by D'Agostino's test [24], mesokurtosis by Anscombe-Glynn test [40], and normality by the Anderson-Darling [41], Shapiro-Francia [42-43], and D'Agostino-Belanger-D'Agostino [44] tests. The difference between asymptotic and bootstrap standard errors was evaluated using the one-sample chi-square test [45].

To facilitate the use of these measures of asymmetry, a script was developed for the R program, allowing for both point and interval estimation. The script adjusts the results based on the sampling distributions of the three indexes studied. Both asymptotic and bootstrap confidence intervals are computed, enabling the testing of the normality of the random sample of the quantitative variable X to determine which interval to use. Three methods are included to obtain the bootstrap confidence intervals: normal, percentile, and bias-corrected and accelerated percentile. Thus, the symmetry (D'Agostino test [24]), mesokurtosis (Anscombe-Glynn test [40]), and normality (Shapiro-Francia [42-43] and D'Agostino-Belanger-D'Agostino [44] tests) of the bootstrap sampling distribution are tested, and the bias and acceleration (a) are reported to select the appropriate method. As an example, the script was applied to a randomly generated sample of 60 data points. To make the example more meaningful, the sample data were generated to correspond to the proportional (monthly) return of stock prices in a volatile stock market, using a Cauchy distribution with a location parameter (x_0) of 0.05, representing the population median, and a scale parameter (γ) of 0.38, constituting the population semi-interquartile range [46-47].

5. Results

5.1. Sampling distribution of *AISr*, *RISr* and *PCS*

The random samples of 10,000 data points (source samples) are plotted as histograms (50 bins of uniform width) with the curve of the theoretical or generating density function overlaid. Figure 1 shows the sample randomly drawn from the standard Wigner semicircular distribution (unit radius). Figure 2 displays the sample randomly drawn from the standard normal distribution, and Figure 3 shows the sample randomly drawn from the standard logistic distribution. From these three source samples, 1,000 samples of 10,000 data points each were drawn with replacement. In each of these samples, Kelley's absolute and relative reversed indices, as well as the percentile coefficient of skewness, were calculated, thus obtaining the bootstrap sampling distribution of these measures of skewness.

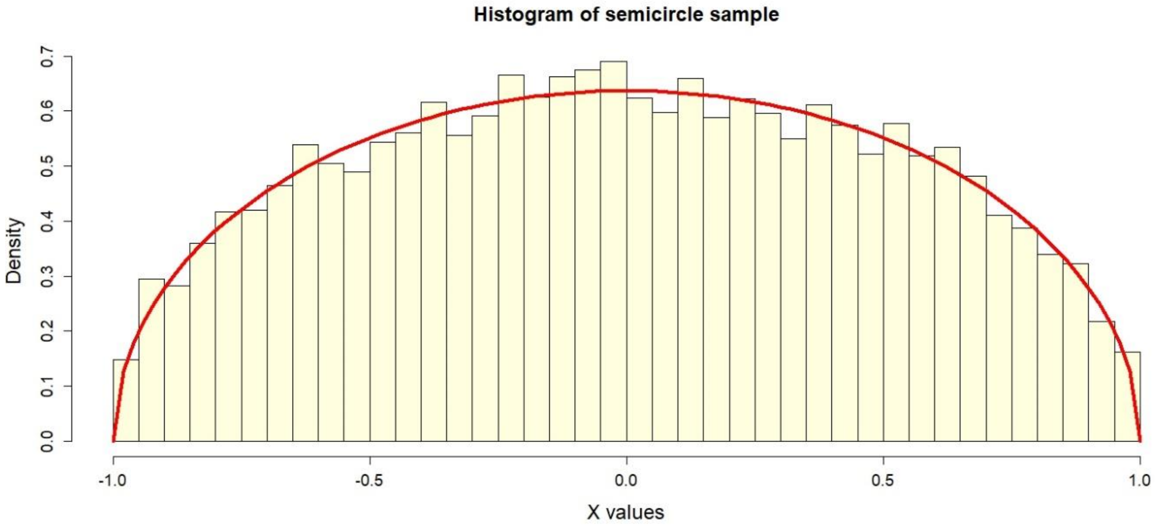

Figure 1. Histogram with the overlaid standard semicircular distribution curve. Source: prepared by the author.

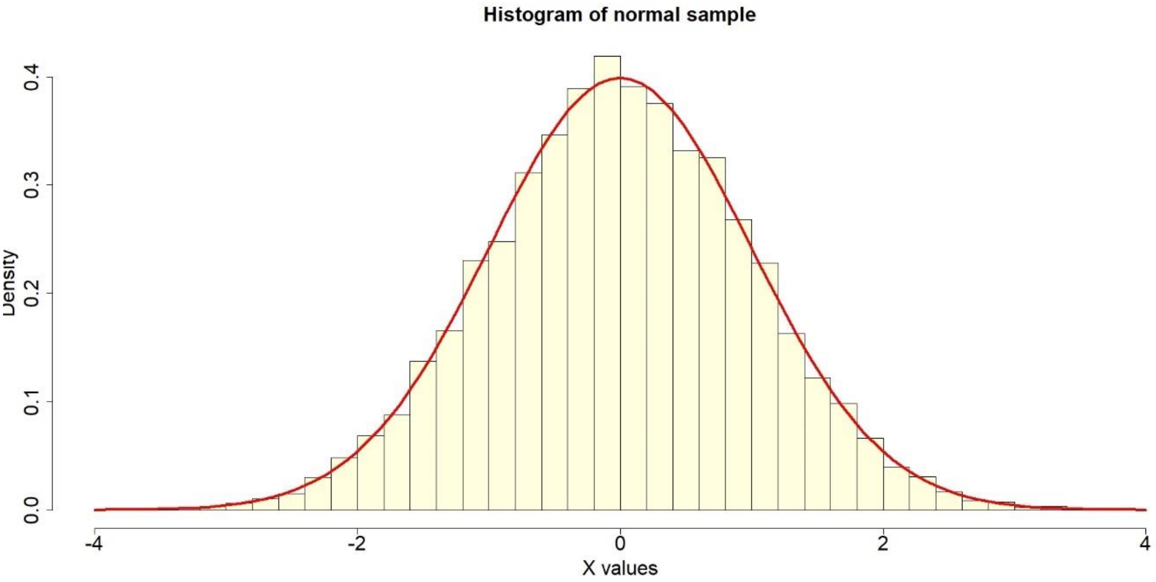

Figure 2. Histogram with the overlaid standard normal curve. Source: prepared by the author.

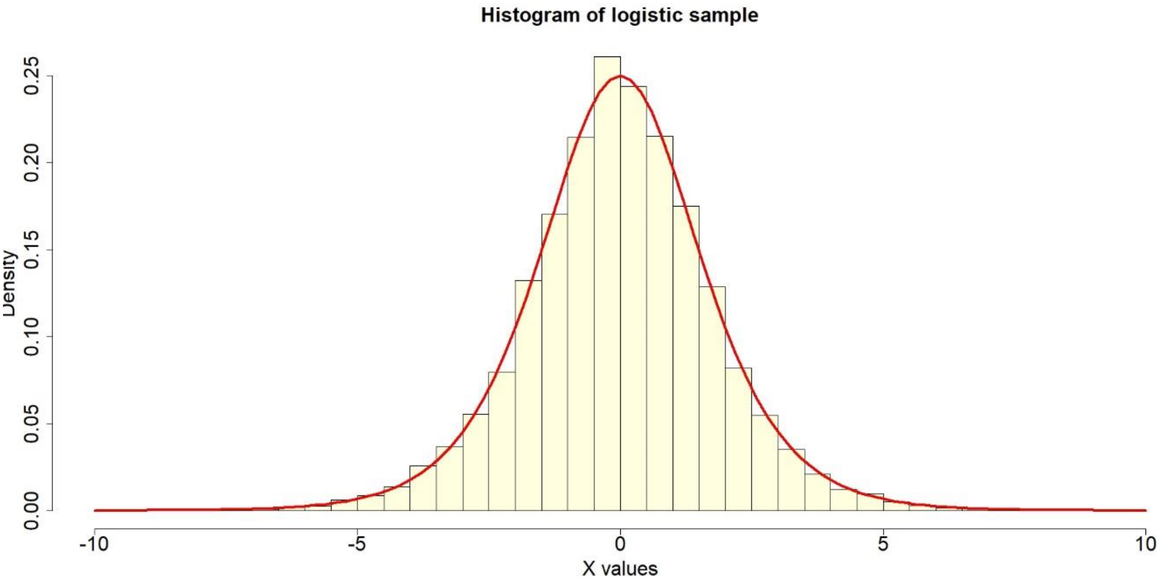

Figure 3. Histogram with the overlaid standard logistic curve. Source: prepared by the author.

At the 5% significance level, the bootstrap sampling distributions of *AISr* showed no outliers by Grubb's test in the three samples generated from 10,000 data points from three symmetric distributions with different kurtosis. Their profiles were symmetric by the D'Agostino test and mesokurtic by the Anscombe-Glynn test. They were fitted to a normal distribution by the Anderson-Darling, Shapiro-Francia, and D'Agostino-Pearson tests. Thus, the asymptotic and bootstrap standard errors were very similar, with absolute differences less than or equal to 0.0010, resulting in a non-significant difference in the bootstrap sampling distribution derived from the logistic distribution (Tables 1-3). Bootstrap standard errors were small, ranging from 0.007 to 0.026. In the normal Q-Q plots, an alignment of the points around the center line at 45 degrees can be seen, and in the histogram, the bell-shaped profile of the normal distribution can be observed, whether the bootstrap sampling distribution is derived from semicircular, normal, or logistically distributed data (Figures 4-6).

The same results of good fit to normality were also observed in the bootstrap sampling distribution of *PCS*, with absolute differences between bootstrap and asymptotic standard errors less than or equal to 0.0015, resulting in a non-significant difference in the bootstrap sampling distribution derived from the sample generated from the logistic distribution. Bootstrap standard errors were small, ranging from 0.011 to 0.012. See Tables 1-3 and Figures 7-9.

However, the bootstrap sampling distributions of *RISr* deviated from normality, presenting outliers, skewness (positive with data from the semicircular and logistic distributions and negative with data from the normal distribution), and marked leptokurtosis (Tables 1-3). The profile of the normal Q-Q plot shows central flattening with a concave curvature at the left end and convex at the right end, which is typical of a strongly leptokurtic distribution. This non-normal, pointed profile can be observed in the histogram, whether the bootstrap sampling distribution is obtained from semicircular, normal, or logistically distributed data (Figures 10-11). Bootstrap standard errors were very large. The largest occurred with the mesokurtic distribution (*bse* = 38.998), followed by the leptokurtic (*bse* = 13.294) and platykurtic (*bse* = 4.897) distributions. Consequently, these estimates are not reliable.

Table 1. Tests of randomness, symmetry and meso-kurtosis for the bootstrap sampling distribution of three measures of skewness.

Type of distribution	Statistic of skewness	Grubbs' test			D'Agostino's test			Anscombe-Glynn test		
		g	u	*p*-value	$\sqrt{b_1}$	z	*p*-value	b_2	z	*p*-value
Semi-circle	AISr	3.307	0.989	0.457	-0.060	-0.774	0.439	2.884	-0.687	0.492
	RISr	29.773	0.112	<0.001	26.136	44.543	<0.001	788.660	23.624	<0.001
	PCS	3.302	0.990	0.467	-0.058	-0.748	0.454	2.887	-0.669	0.504
Normal	AISr	2.984	0.991	0.999	0.035	0.456	0.649	2.900	-0.571	0.568
	RISr	29.376	0.135	<0.001	-25.454	-44.260	<0.001	747.730	23.566	<0.001
	PCS	3.027	0.991	0.999	0.0247	0.321	0.749	2.905	-0.538	0.591
Logistic	AISr	3.596	0.987	0.155	-0.040	-0.519	0.604	3.014	0.203	0.839
	RISr	18.725	0.649	<0.001	6.065	28.904	<0.001	168.091	21.445	<0.001
	PCS	3.490,	0.988	0.233	-0.032	-0.417	0.677	2.999	0.106	0.915

[1] Note. Statistic: *AISr* = Reversed Absolute Index of Skewness, *RISr* = Reversed Relative Index of Skewness, and *PCS* = Percentile Coefficient of Skewness. Grubbs' test for one outlier: *g* = Grubbs' test statistic, *u* = normalized test statistic, and *p*-value = one-tailed p-value (alternative hypothesis: lowest or highest value is an outlier). D'Agostino skewness test: $\sqrt{b_1}$ = measure of skewness based on the standardized third central moment, *z* = test statistic, and p-value = two-tailed *p*-value (alternative hypothesis: data have a skewness towards left or right tail). Anscombe-Glynn kurtosis test: b_2 = measure of kurtosis based on the standardized fourth central moment. *z* = test statistic, and *p*-value = two-tailed *p*-value (alternative hypothesis: the value of kurtosis [β_2] is not equal to 3). Probability values less than 0.05, indicating the statistical test's significance, are highlighted in bold. Source: prepared by the author.

Table 2. Tests of normality for the bootstrap sampling distribution of three measures of skewness.

Type of distribution	**Statistic of skewness**	**Anderson-Darling**		**Shapiro-Francia**		**D'Agostino et al.**	
		AD	***p*-value**	**w**	***p*-value**	**k^2**	***p*-value**
Semi-circle	AISr	0.172	0.930	0.999	0.830	1.072	0.585
	RISr	285.900	< 0.001	0.088	< 0.001	2542.163	< 0.001
	PCS	0.167	0.938	0.999	0.863	1.007	0.604
Normal	AISr	0.312	0.551	0.999	0.631	0.533	0.766
	RISr	267.320	< 0.001	0.110	< 0.001	2514.260	< 0.001
	PCS	0.319	0.534	0.999	0.675	0.392	0.822
Logistic	AISr	0.288	0.617	0.998	0.434	0.311	0.856
	RISr	215.960	< 0.001	0.284	< 0.001	1295.340	< 0.001
	PCS	0.2873	0.620	0.998	0.475	0.185	0.912

[1] Note. Statistic: *AISr* = Reversed Absolute Index of Skewness, *RISr* = Reversed Relative Index of Skewness, and *PCS* = Percentile Coefficient of Skewness Anderson-Darling normality test: *AD* = test statistic, and *p*-value (alternative hypothesis: sample was not drawn from a normal distribution). Shapiro-Francia normality test: *w* = test statistic, and *p*-value (the same alternative hypothesis). D'Agostino-Berlanger-D'Agostino Normality test: k^2 = test statistic, and *p*-value (the same alternative hypothesis). Probability values less than 0.05, indicating the statistical test's significance, are highlighted in bold. Source: prepared by the author.

Table 3. Point estimation of skewness measures and bootstrap confidence interval, bias and comparison between bootstrap and asymptotic standard errors.

Distribution	Statistic	value	Method	LL	UL	bias	a	bse	ase	diff	*p*-value
Semi-circle	AISr	0.004	norm	-0.010	0.018	0.00002	0.00043	0.007	0.008	-0.0010	**<0.001**
	RISr	-0.507	BCa	-2.029	4.113	-0.0603	-0.00014	4.897	1.031	3.8669	**<0.001**
	PCS	0.006	norm	-0.015	0.026	0.00003	0.00044	0.011	0.012	-0.0015	**<0.001**
Normal	AISr	0.003	norm	-0.023	0.029	-0.0004	0.00163	0.013	0.015	-0.0018	**<0.001**
	RISr	-0.255	BCa	-3.759	12.332	-0.7699	-0.00212	38.998	1.374	37.6246	**<0.001**
	PCS	0.002	norm	-0.018	0.023	-0.0003	0.00162	0.011	0.012	-0.0015	**<0.001**
Logistic	AIS	-0.004	norm	-0.055	0.045	0.0011	0.00007	0.026	0.026	-0.0003	0.613
	RIS	0.169	BCa	-5.167	21.436	-0.1044	-0.00004	13.294	1.190	12.1045	**<0.001**
	PCS	-0.002	norm	-0.025	0.021	0.0005	0.00009	0.012	0.012	-0.0002	0.461

[1] Note. Statistic: *AISr* = Reversed Absolute Index of Skewness, *RISr* = Reversed Relative Index of Skewness, and *PCS* = Percentile Coefficient of Skewness. Method: norm = normal from library(boot) and BCa = bias-corrected and accelerated bootstrap from library 'bcajack'. *LL* = low limit and *UL* = upper limit of the 95% bootstrap confidence interval, bias = difference between the bootstrap estimate (mean in the bootstrap sampling distribution of the statistic) and the original sample estimate, *bse* = bootstrap standard error, *ase* = asymptotic standard error, *diff* = difference between errors: *bse* - *ase*, *p*-value = one-tailed *p*-value for chi-square test of a single variance (alternative hypothesis: bootstrap variance is lower (left tail) or higher (right tail) than the asymptotic variance expected from the normal approximation). Probability values less than 0.05, indicating the statistical test's significance, are highlighted in bold. Source: prepared by the author.

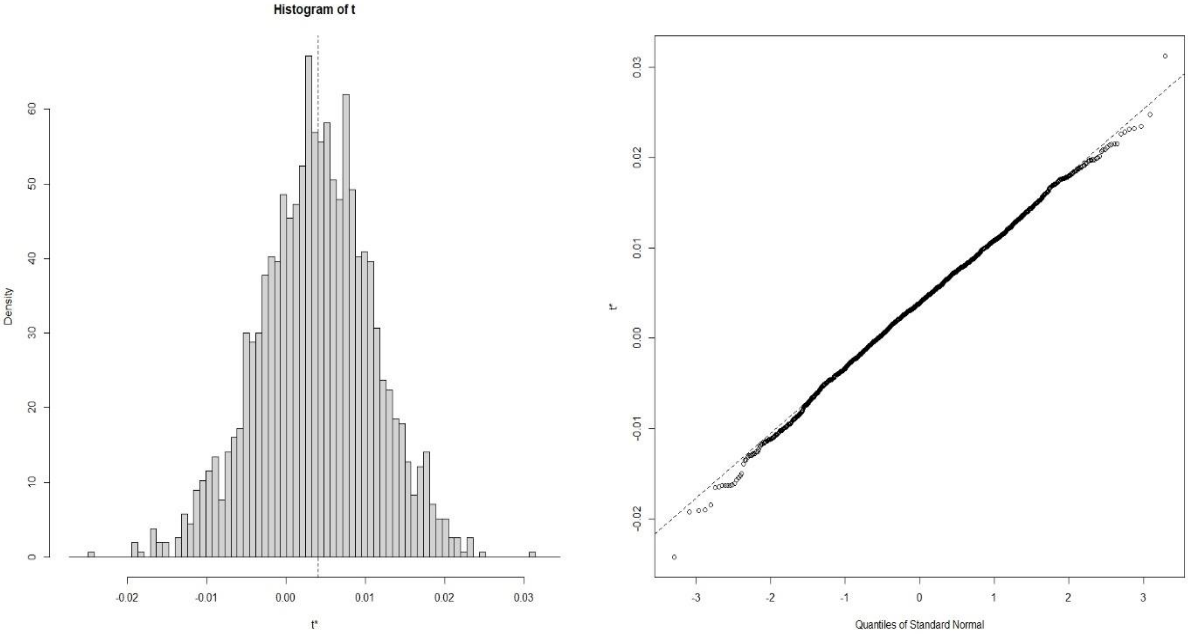

Figure 4. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *AISr* from the 10,000-data sample drawn from a semicircle distribution. Source: prepared by the author.

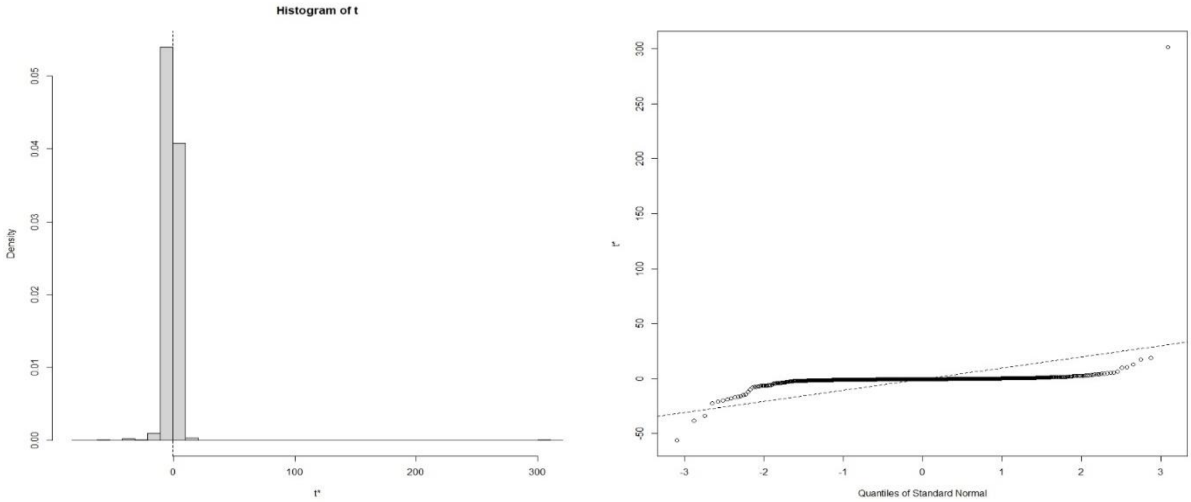

Figure 5. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *RISr* from the 10,000-data sample drawn from a semicircle distribution. Source: prepared by the author.

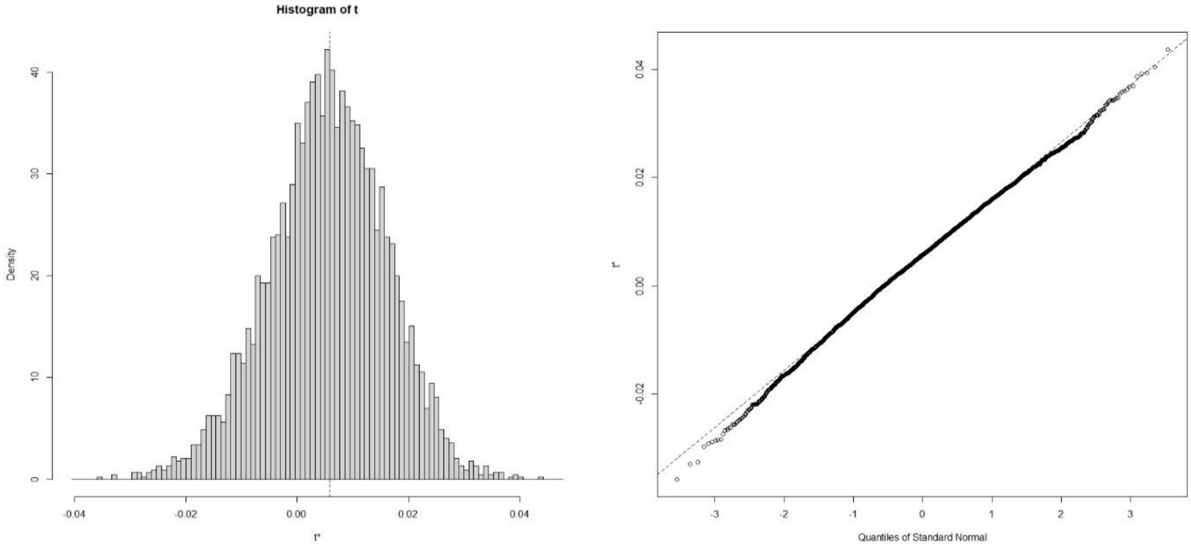

Figure 6. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *PCS* from the 10,000-data sample drawn from a semicircle distribution. Source: prepared by the author.

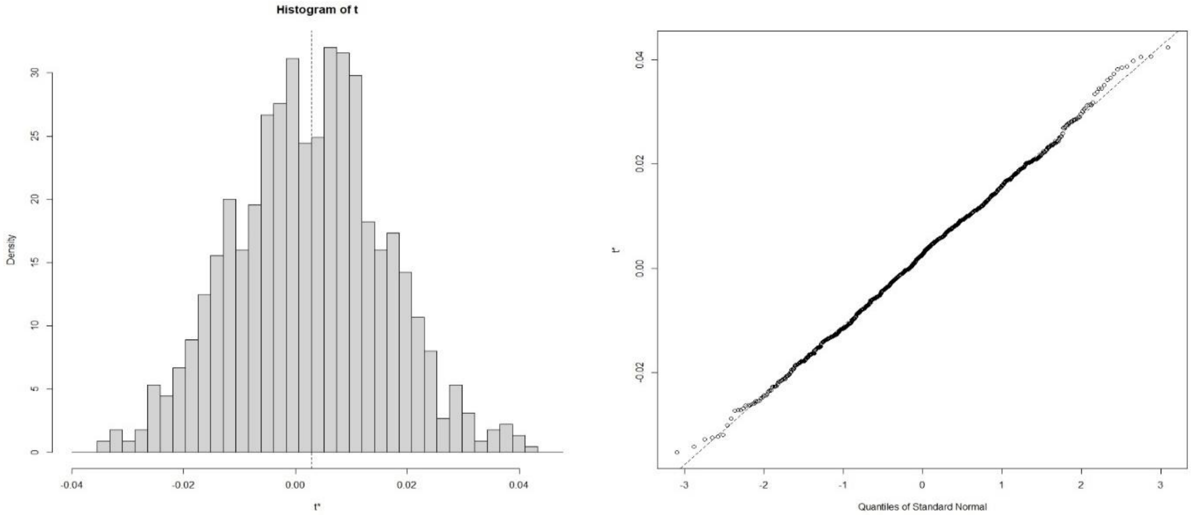

Figure 7. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *AISr* from the 10,000-data sample drawn from a normal distribution. Source: prepared by the author.

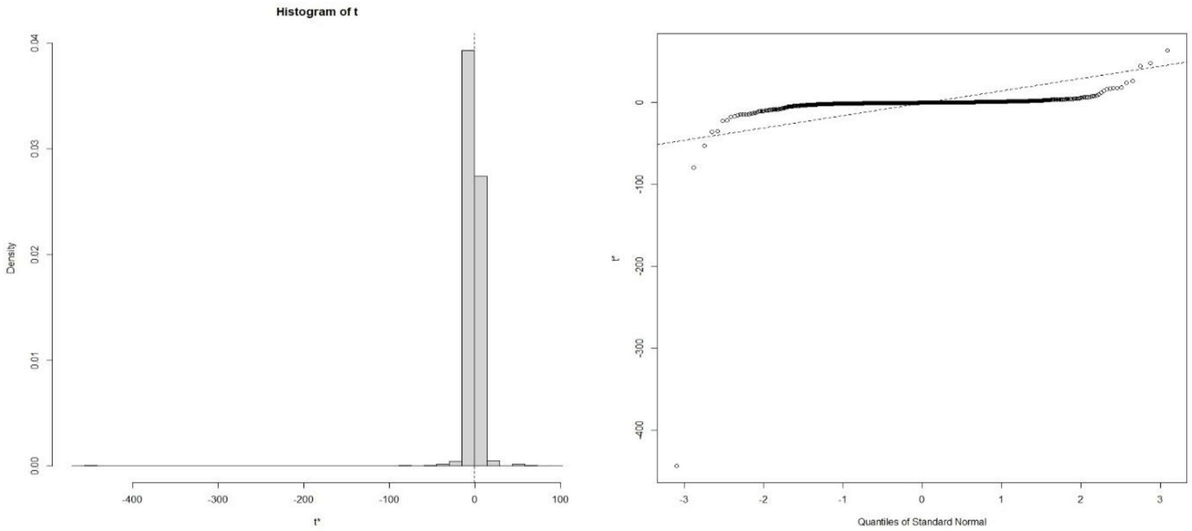

Figure 8. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *RISr* from the 10,000-data sample drawn from a normal distribution. Source: prepared by the author.

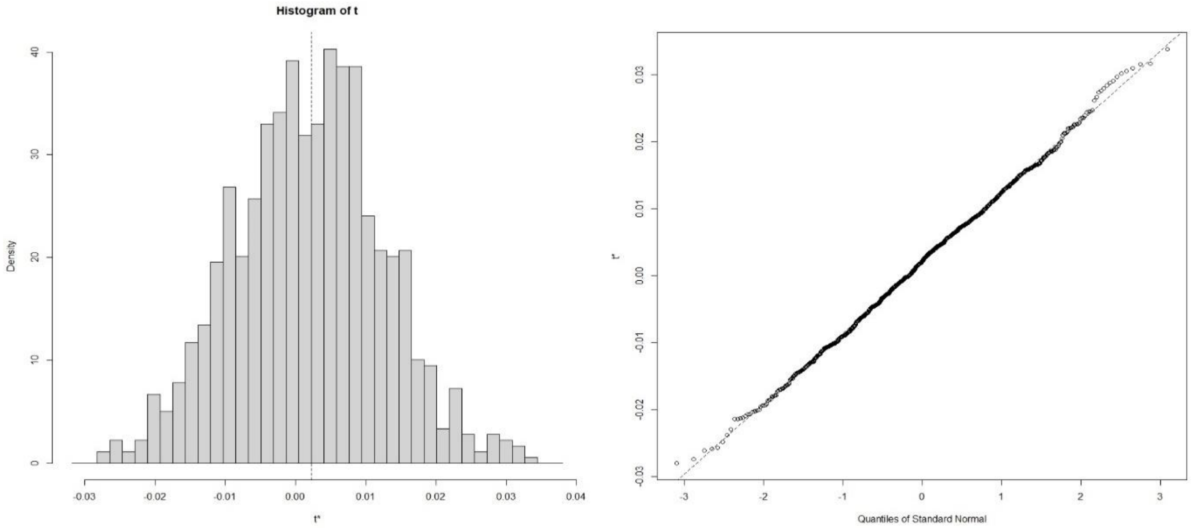

Figure 9. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *PCS* from the 10,000-data sample drawn from a normal distribution. Source: prepared by the author.

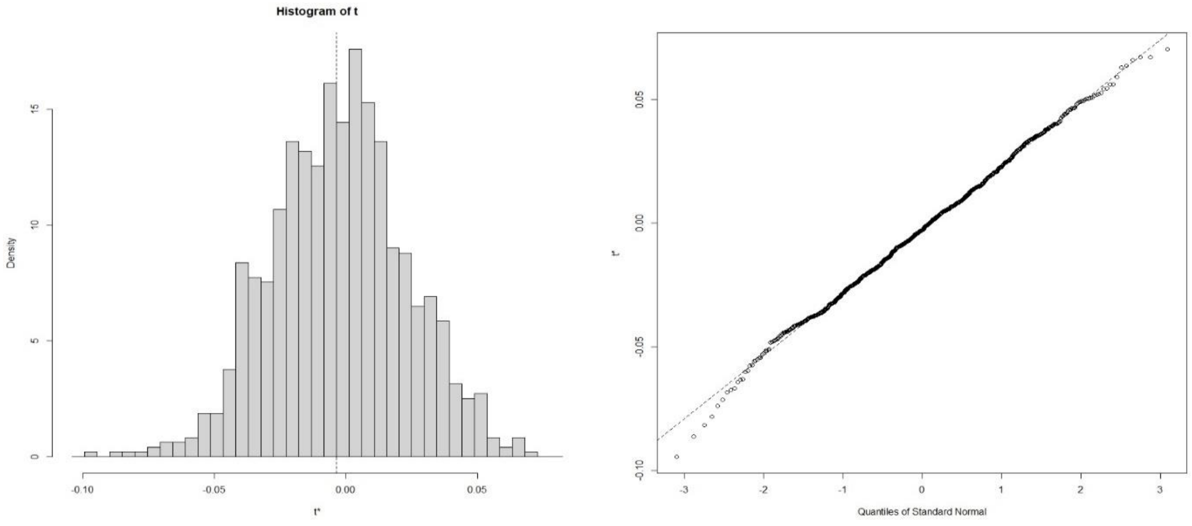

Figure 10. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *AISr* from the 10,000-data sample drawn from a logistic distribution. Source: prepared by the author.

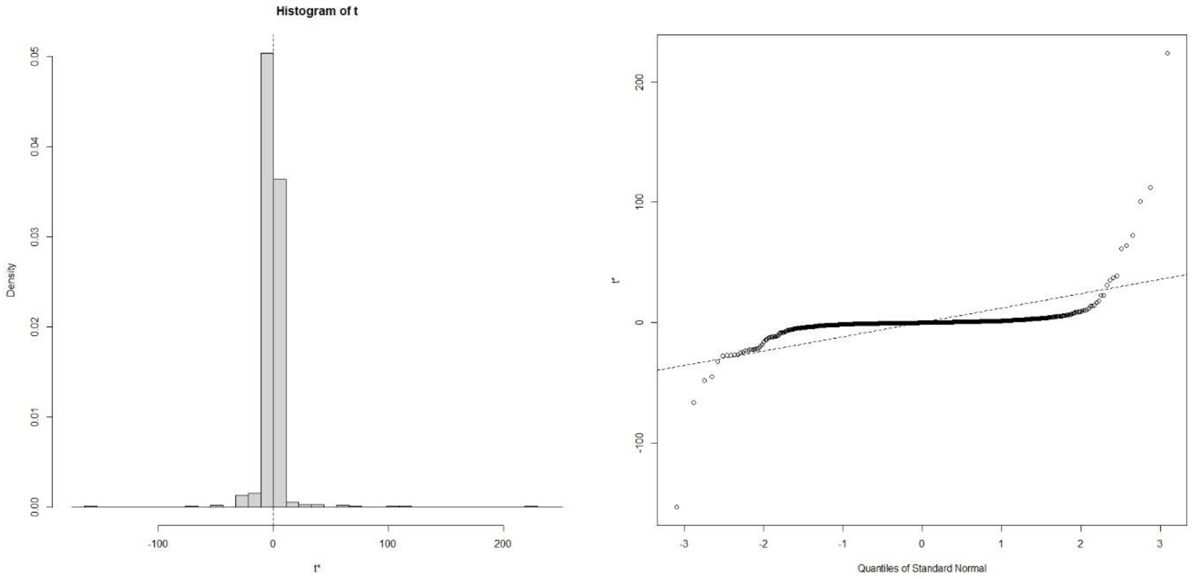

Figure 11. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *RISr* from the 10,000-data sample drawn from a logistic distribution. Source: prepared by the author.

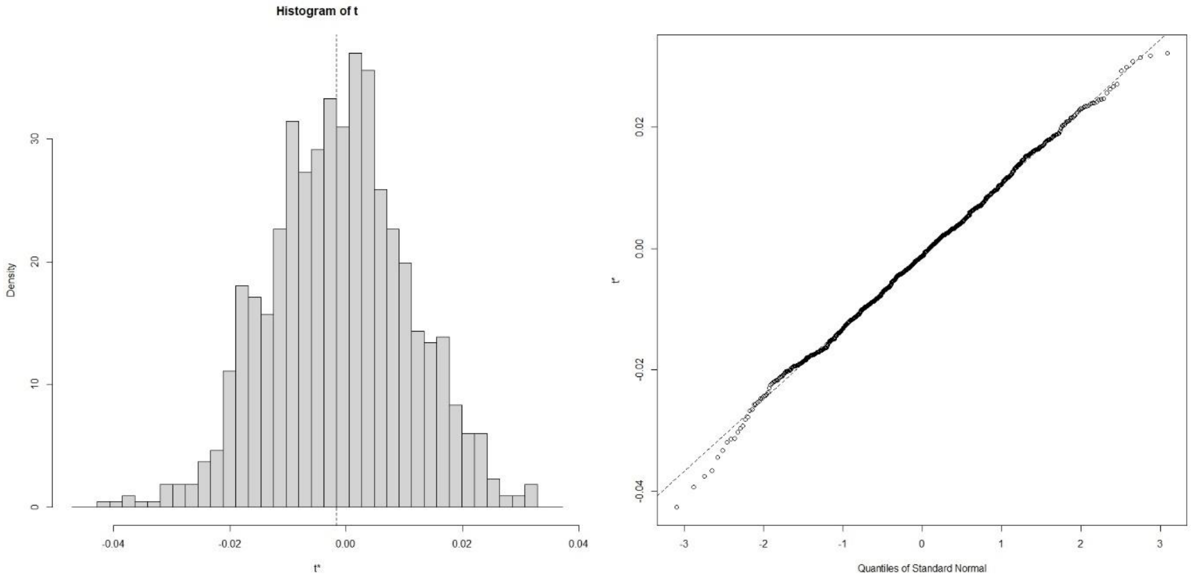

Figure 12. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution of *PCS* from the 10,000-data sample drawn from a logistic distribution. Source: prepared by the author.

5.2. Script for point and interval estimation of *AISr* and *PCS*

5.2.1. Content and structure of the script

The randomness of the sample is a fundamental assumption for statistical inference [48], so the script starts by testing this assumption using the Wald-Wolfowitz runs test. Running the test in the R program requires downloading the 'randtests' library [49].

For this and the other inferential tests in the script, the significance level is set at 0.05, which is the most conventional value [50]. However, it can be modified considering the sample size [51]. It can be raised to 0.1 with small samples (20 to 29) or lowered to 0.01 with large samples (1000 or more). It can also be changed based on a theoretical approach [52] or to control for family-wise error rate [53].

The a posteriori statistical power (Φ) is calculated by bootstrapping. From the original sample, 1000 samples are generated, and for each of these 1000 samples, the exact probability of the two-tailed test (p-value = 'exact') is calculated. However, one can switch to the asymptotic probability in the case of a large sample, where n_0 (the number of data points smaller than the median or criterion) and n_1 (the number of data points larger than the median) are each greater than or equal to 20 (p-value = 'normal'). The number of probabilities less than the significance level, divided by the number of simulations, yields the statistical power or the probability of rejecting the null hypothesis when it is false [26]. See Equation 13, where I is the indicator function, B represents the number of simulations, and α denotes the significance level.

$$\phi = \frac{\sum_{i=}^{B} I\left(p_{two-sided\_exat} < \alpha\right)}{B} \tag{13}$$

The script continues by testing for normality. In case of a good fit, the use of the asymptotic error and confidence interval for $AISr$ and PCS can be chosen. This distributional assumption is tested through four tests with different inferential foundations [54], and additionally, the statistical power is computed. Executing these tests in R requires loading the 'moments' [55] and 'nortest' [56] libraries.

The normality check starts with the D'Agostino-Berlanger-D'Agostino K^2 test [44] based on the sum of squares of the standardized values of skewness [24] and kurtosis [40], whose statistic converges in distribution to a chi-square distribution with two degrees of freedom. With this test, a posteriori statistical power is calculated with the complementary cumulative distribution function of a non-central chi-square distribution.

The second normality test is the Lilliefors test [57], which assesses the maximum linear distance between the empirical and theoretical cumulative distribution functions. The third test is the Anderson-Darling test [41], which measures the standardized quadratic distance between these functions. The fourth and final test is the Shapiro-Francia test [42-43], based on the shared variance between empirical and theoretical quantiles. With these three tests, a posteriori statistical power is obtained through bootstrapping. They were chosen for their recommended power and adequacy [58-61].

The assessment of normality through inferential tests is complemented by graphical representation using a histogram with overlaid density and normal curves. The uniform amplitude and number of bins are determined using the Freedman-Diaconis optimization rule (breaks = "fd"), which imposes no distributional assumption. However, this can be switched to the Scott's rule (breaks = "scott") if the data fit well to a normal distribution [62]. The density for the overlaid curve is computed using the Epanechnikov's kernel function [63], known for its parabolic shape that minimizes integrated mean square error most effectively [64]. The bandwidth is computed using the Sheather-Jones procedure [65], which minimizes asymptotic integrated mean square error and is considered one of the most suitable methods for this purpose [66-67].

It should be noted that when running the script using the R program downloaded on your computer, the plot can be saved as a high-definition JPEG file. Simply expand the R graphics device window (ACTIVE), where in its toolbar under Files, you will find the 'Save As' function that offers this option [68].

After these checks, the script proceeds to the point and interval estimation of the additive complement of AIS and PCS. The additive complement of the IAS was chosen for two reasons: firstly, to maintain interpretative consistency with the classic index established by Pearson [3-4] — positive values indicate right-tailed skewness, and negative values indicate left-tailed skewness. Secondly, because PCS represents the standardized form of the additive complement of AIS once it is divided by the semi-percentile range. Quantiles were calculated using R's rule 8 [18-19]. $RISr$ is omitted from the script due to its excessive bootstrap standard error. However, its calculation with R applied to a sample of 60 data points is shown at the end of this Results section (sub-section 5.3).

The confidence level is set at 95%, corresponding to the 5% significance level chosen for the inferential tests. For the bootstrap confidence intervals, 1000 random samples were created by sampling with replacement from the original sample, and the statistics $AISr$ and PCS were calculated for each of these 1000 samples [26]. Thus, bootstrap sampling distributions were generated for these two statistics. The symmetry of the bootstrap sampling distribution was assessed using the D'Agostino test [24], kurtosis was tested using the Anscombe-Glynn test [40], and normality was evaluated using the Shapiro-Francia test [42-43]. Additionally, bootstrap bias, bootstrap standard error (*bse*), jackknife acceleration (a), and the density histogram and Q-Q normal plot of the bootstrap sampling distributions were obtained [23].

Bootstrap confidence intervals are calculated using three methods: normal (norm), percentile (perc), and bias-corrected and accelerated percentile (BCa). The choice of method depends on the characteristics of the bootstrap sampling distribution. The normal method is preferred when the distribution is approximately normal. If there is no normality, bias ($|b| \geq 0.05$) and asymmetry ($|a| \geq 0.025$ and significance D'Agostino's skewness test [24]), then the BCa method is the best choice [25]. Alternatively, when there is no normality, but acceleration and bias are minimal, the percentile method remains suitable. To implement these methods, downloading the 'boot' library is necessary [69].

Finally, the bootstrap probability of retaining the null hypothesis of symmetry, assessed using *AISr* and *PCS*, was calculated. This is obtained by counting how many values are less than or equal (left tail) and how many values are greater than or equal (right tail) to the parameter (null hypothesis of symmetry) among the 1000 data from the statistic's bootstrap sampling distribution of *AISr* or *PCS*. The smallest probability is chosen and multiplied by 2. If it is less than or equal to the significance level (α), the null hypothesis of symmetry holds in a two-tailed test, otherwise it is rejected [70].

A random sample of at least 20 data points [71] and preferably at least 30 data points [72] from a quantitative variable X is recommended for using this script. The script is divided into three independent parts to make it lighter to execute. What can be changed in the script to match the user's data is marked in blue, such as the sample or vector of scores.

5.2.2. The script with its three parts

```
# First part of the script: randomness and normality of the sample or vector of scores x.
x <- c(0.532, -0.248, 1.342, -0.096, -0.022, 0.105, -4.243, -0.084, -2.281, 2.821, -0.002, 2.623, -0.559, -1.586,
0.177, -0.074, 0.421, 0.101, 0.683, -0.005, -0.087, -0.499, -0.754, 0.043, -0.664, -0.445, -2.677, -1.197, 0.537,
0.239, 0.006, -0.070, -0.506, -0.234, 0.079, 5.489, -0.310, 0.357, 0.641, 0.388, 0.233, 1.431, 1.417, 0.919, 0.247,
0.227, 0.392, 3.590, 0.470, -0.132, 0.105, 2.120, -0.228, 0.203, -1.910, 0.338, 0.211, -0.322, -0.080, 0.963)

cat("Testing the randomness of the sample x using the Wald-Wolfowitz runs test", "\n")
library(randtests)
runse <- runs.test(x, alternative = "two.sided", threshold = median(x), pvalue = 'exact')
runsa <- runs.test(x, alternative = "two.sided", threshold = median(x), pvalue = 'normal')
alpha <- 0.05
ww_power <- function(x, alpha, B = 1000) {n <- length(x)
p_values <- numeric(B)
for (i in 1:B) {bootstrap_sample <- sample(x, replace = TRUE)
result <- runs.test(x, alternative = "two.sided", threshold = median(x), pvalue = 'exact')
p_values[i] <- result$p.value}
power <- mean(p_values < alpha)
return(power)}
set.seed(123)
power <- ww_power(x, alpha)
cat("Wald-Wolfowitz runs test. Criterion: median", "\n")
cat("Number of runs: r =", runse$runs, "\n")
cat("n_0 = #(x_i < mdn(x)) =", runse$parameter["n1"],"y", "n_1 = #(x_i > mdn(x)) =",
runse$parameter["n2"], "\n")
cat("n = n_0 + n_1 =", runse$parameter["n"], "\n")
cat("Two-tailed exact probability value: p =", round(runse$p.value, 3), "\n")
cat("Mean: M(R|n_0, n_1) =", runse$mu, "y", " Standard deviation: DE(R|n_0, n_1) =",
round(sqrt(runse$var), 3), "\n")
cat("Standardized number of runs: z_r =", round(runse$statistic, 3), "\n")
cat("Two-tailed asymptotic probability value: p =", round(runsa$p.value, 3), "\n")
cat("Statistical power for the Wald-Wolfowitz runs test using bootstrap simulation: ϕ =", power, "\n")
cat("Testing for normality using four tests with different rationales:", "\n")
library(moments)
```

```
agostino <- agostino.test(x, alternative = "two.sided")
ag <- anscombe.test(x, alternative = "two.sided")
K2 <- (agostino$statistic["z"]^2 + ag$statistic["z"]^2)
p <- pchisq(K2, df = 2, lower.tail = FALSE)
power_K2 <- 1 - pchisq(qchisq(alpha, df = 2, lower.tail = FALSE), df = 2, ncp = K2, lower.tail = TRUE, log.p = FALSE)
cat("Test for normality using the D'Agostino-Berlanger-D'Agostino test.", "\n")
cat("Test statistic: k^2(x) =", round(K2, 3), "\n")
cat("Asymptotic probability value in a two-tailed test: p_value =", round(p, 3), "\n")
if (p < alpha) {cat(sprintf("The null hypothesis of normality is rejected at a significance level of %.2f using K squared test.", alpha), "\n")
} else {cat(sprintf("The null hypothesis of normality is maintained at a significance level of %.2f using K squared test.", alpha), "\n")}
cat("The right-tailed statistical power for the alternative hypothesis of non-normality for the K squared test: ϕ =", round(power_K2, 4), "\n")
library(nortest)
lillie.test(x)
result_lillie <- lillie.test(x)
if (result_lillie$p.value < alpha) {cat(sprintf("The null hypothesis of normality is rejected at a significance level of %.2f using Lilliefors test.", alpha), "\n")
} else {cat(sprintf("The null hypothesis of normality is maintained at a significance level of %.2f using Lilliefors test.", alpha), "\n")}
lillie_power <- function(x, alpha, B = 1000) {n <- length(x)
p_values <- numeric(B)
for (i in 1:B) {bootstrap_sample <- sample(x, replace = TRUE)
result <- lillie.test(bootstrap_sample)
p_values[i] <- result$p.value}
power <- mean(p_values < alpha)
return(power)}
set.seed(123)
power <- lillie_power(x, alpha)
cat("Statistical power for the Lilliefors normality test using bootstrap simulation: ϕ =", power, "\n")

ad.test(x)
result_ad <- ad.test(x)
if (result_ad$p.value < alpha) {cat(sprintf("The null hypothesis of normality is rejected at a significance level of %.2f using Anderson-Darling test.", alpha), "\n")
} else {cat(sprintf("The null hypothesis of normality is maintained at a significance level of %.2f using Anderson-Darling test.", alpha), "\n")}
ad_power <- function(x, alpha, B = 1000) {n <- length(x)
p_values <- numeric(B)
for (i in 1:B) {bootstrap_sample <- sample(x, replace = TRUE)
result <- ad.test(bootstrap_sample)
p_values[i] <- result$p.value}
power <- mean(p_values < alpha)
return(power)}
set.seed(123)
power <- ad_power(x, alpha)
cat("Statistical power for the Anderson-Darling normality test using bootstrap simulation: ϕ =", power, "\n")
```

```
sf.test(x)
result_sf <- sf.test(x)
if (result_sf$p.value < alpha) {cat(sprintf("The null hypothesis of normality is rejected at a significance level of %.2f using Shapiro-Francia test.", alpha), "\n")
} else {cat(sprintf("The null hypothesis of normality is maintained at a significance level of %.2f using Shapiro-Francia test.", alpha), "\n")}
sf_power <- function(x, alpha, B = 1000) {n <- length(x)
p_values <- numeric(B)
for (i in 1:B) {bootstrap_sample <- sample(x, replace = TRUE)
result <- sf.test(bootstrap_sample)
p_values[i] <- result$p.value}
power <- mean(p_values < alpha)
return(power)}
set.seed(123)
power <- sf_power(x, alpha)
cat("Statistical power for the Shapiro-Francia normality test using bootstrap simulation: ϕ =", power, "\n")

# Histogram with overlaid density and normal curves
# breaks = "scott" in case of normality or breaks = "sturges" in the case of symmetry
hist(x, breaks = "fd", col = "darkolivegreen2", border = "black", freq = FALSE, main = "Histogram with overlaid normal and density curves", xlab = " X values", ylab = "Density", ylim = c(0, 0.8), cex.main = 1.5, cex.lab = 1.5, cex.axis = 1.2)
lines(density(x, kernel = "epanechnikov", bw = "sj"), col = "darkblue", lwd = 4)
x_values <- seq(mean(x) - 4 * sd(x), mean(x) + 4 * sd(x), length = 1000)
y_values <- dnorm(x_values, mean = mean(x), sd = sd(x))
lines(x_values, y_values, col = "red", lwd = 4)

# Second part of the script: AISr
x <- c(0.532, -0.248, 1.342, -0.096, -0.022, 0.105, -4.243, -0.084, -2.281, 2.821, -0.002, 2.623, -0.559, -1.586, 0.177, -0.074, 0.421, 0.101, 0.683, -0.005, -0.087, -0.499, -0.754, 0.043, -0.664, -0.445, -2.677, -1.197, 0.537, 0.239, 0.006, -0.070, -0.506, -0.234, 0.079, 5.489, -0.310, 0.357, 0.641, 0.388, 0.233, 1.431, 1.417, 0.919, 0.247, 0.227, 0.392, 3.590, 0.470, -0.132, 0.105, 2.120, -0.228, 0.203, -1.910, 0.338, 0.211, -0.322, -0.080, 0.963)

library(boot)
library(moments)
library(nortest)
alpha <- 0.05

cat("Point estimate of the additive opposite of Kelley's absolute skewness index", "\n")
n <- length(x)
q0.1 <- quantile(x, 0.10, type = 8)
q0.5 <- quantile(x, 0.50, type = 8)
q0.9 <- quantile(x, 0.90, type = 8)
AISr = (q0.1 + q0.9) / 2 - q0.5
cat("Sample size: n =", n,"\n")
cat("Quantile of order 0.1 by the rule 8: q_0.1(x) =",round(q0.1, 3),"\n")
cat("Quantile of order 0.5 by the rule 8: q_0.5(x) =",round(q0.5, 3),"\n")
cat("Quantile of order 0.1 by the rule 8: q_0.9(x) =",round(q0.9, 3),"\n")
cat("Additive opposite of Kelley's absolute asymmetry index: AISr = ",round(AISr, 3),"\n")
```

```
cat("Asymptotic confidence interval of the additive opposite of Kelley's absolute skewness index", "\n")
ase <- 0.599143 * (quantile(x, 0.9, type = 8) - quantile(x, 0.10, type = 8)) / sqrt(length(x))
LL = AISr - qnorm(1-alpha/2) * ase
UL = AISr + qnorm(1-alpha/2) * ase
cat("Asymptotic standard error of AISr: ase(AISr) =", round(ase, 3), "\n")
cat("Asymptotic confidence interval", (1 - alpha) * 100, "% for AISr: 95% CI (", round(LL, 3),",", round(UL, 3),")\n")

# Acceleration (jackknife)
AISr_jack <- numeric(n)
for (i in 1:n) {x_jack <- x[-i]
AISr_jack[i] <- quantile((quantile(x_jack, 0.10, type = 8) + quantile(x_jack, 0.90, type = 8))/2 - x_jack, 0.50, type = 8)}
acel <- sum((mean(AISr_jack) - AISr_jack)^3) / (6 * sum((mean(AISr_jack) - AISr_jack)^2)^(3/2))
cat("Aceleración (jackknife): a =", round(acel, 6), "\n")
set.seed(123)
b <- boot(data = x, function(x, i) {(quantile(x[i], 0.10, type = 8) + quantile(x[i], 0.90, type = 8))/2 - quantile(x[i], 0.50, type = 8)}, R=1000)
boot_est <- mean(b$t)
cat("Bootstrap estimation: m(b$t) =", round(boot_est, 3), "\n")
agostino_boot <- agostino.test(b$t, alternative = "two.sided")
ag_boot <- anscombe.test(b$t, alternative = "two.sided")
print(agostino_boot)
print(ag_boot)
cat("D'Agostino-Berlanger-D'Agostino normality test", "\n")
K2 <- (agostino_boot$statistic["z"]^2 + ag_boot$statistic["z"]^2)
p <- pchisq(K2, df = 2, lower.tail = FALSE)
cat("K^2 =", round(K2, 3), "\n")
cat("p_value =", round(p, 3), "\n")
sf.test(b$t)
b
plot(b)
boot.ci(b, conf = 0.95, type = c("bca", "perc", "norm"))
p_boot_left <- mean(b$t < 0)
p_boot_right <- mean(b$t > 0)
p_boot <- min(p_boot_left, p_boot_right)
cat("One-tailed bootstrap probability value for the null hypothesis of symmetry: p =", round(p_boot, 3), "\n")
if (p_boot < alpha) {cat(sprintf("The null hypothesis of normality is rejected at a significance level of %.2f using bootstrap probability.", alpha), "\n")
} else {cat(sprintf("The null hypothesis of normality is maintained at a significance level of %.2f using bootstrap probability.", alpha), "\n")}

# Third part of the script: PCS

x <- c(0.532, -0.248, 1.342, -0.096, -0.022, 0.105, -4.243, -0.084, -2.281, 2.821, -0.002, 2.623, -0.559, -1.586, 0.177, -0.074, 0.421, 0.101, 0.683, -0.005, -0.087, -0.499, -0.754, 0.043, -0.664, -0.445, -2.677, -1.197, 0.537, 0.239, 0.006, -0.070, -0.506, -0.234, 0.079, 5.489, -0.310, 0.357, 0.641, 0.388, 0.233, 1.431, 1.417, 0.919, 0.247, 0.227, 0.392, 3.590, 0.470, -0.132, 0.105, 2.120, -0.228, 0.203, -1.910, 0.338, 0.211, -0.322, -0.080, 0.963)

library(bcaboot)
```

```
library(boot)
library(moments)
library(nortest)

cat("Point estimate of the percentile coefficient of skewness", "\n")
PCS <- (quantile(x, 0.10, type = 8) + quantile(x, 0.90, type = 8) - 2* quantile(x, 0.50, type = 8)) / (quantile(x,
0.90, type = 8) - quantile(x, 0.10, type = 8))
cat("Percentage coefficient of skewness: PCS(x) = ", PCS,"\n")

cat("Asymptotic confidence interval of percentile coefficient of skewness", "\n")
ase <- 1.198286 / sqrt(length(x))
alpha <- 0.05
LL <- PCS - qnorm(1 - alpha / 2) * ase
UL <- PCS + qnorm(1- alpha / 2) * ase
cat("Asymptotic standard error of PCS: ase(PCS) =", round(ase, 3), "\n")
cat("Asymptotic confidence interval", (1 - alpha) * 100, "% for PCS: 95% CI (", round(LL, 3),",", round(UL,
3),")\n")

cat("Bootstrap confidence interval of the percentile coefficient of skewness", "\n")
# Aceleración (jackknife)
n <- length(x)
PCS_jack <- numeric(n)
for (i in 1:n) {x_jack <- x[-i]
PCS_jack[i] <- (quantile(x_jack, 0.10, type = 8) + quantile(x_jack, 0.90, type = 8) - 2 * quantile(x_jack, 0.50,
type = 8)) / (quantile(x_jack, 0.90, type = 8) - quantile(x_jack, 0.10, type = 8))}
acel <- sum((mean(PCS_jack) - PCS_jack)^3) / (6 * sum((mean(PCS_jack) - PCS_jack)^2)^(3/2))
cat("Acceleration (jackknife): a =", round(acel, 9), "\n")
library(boot)
set.seed(123)
b<-boot(data=x, function(x, i) {(quantile(x[i], 0.10, type = 8) + quantile(x[i], 0.90, type = 8) - 2*quantile(x[i],
0.50, type = 8))/(quantile(x[i], 0.90, type = 8) - quantile(x[i],0.10, type = 8))}, R=1000)
boot_est <- mean(b$t)
cat("Bootstrap estimation: m(b$t) =", round(boot_est, 3), "\n")
agostino_boot <- agostino.test(b$t, alternative = "two.sided")
ag_boot <- anscombe.test(b$t, alternative = "two.sided")
print(agostino_boot)
print(ag_boot)
cat("D'Agostino-Berlanger-D'Agostino normality test", "\n")
K2 <- (agostino_boot$statistic["z"]^2 + ag_boot$statistic["z"]^2)
p <- pchisq(K2, df = 2, lower.tail = FALSE)
cat("K^2 =", round(K2, 3), "\n")
cat("p_value =", round(p, 3), "\n")
sf.test(b$t)
b
plot(b)
boot.ci(b, conf = 0.95, type=c("bca", "perc", "norm"))
p_boot_left <- mean(b$t < 0)
p_boot_right <- mean(b$t > 0)
p_boot <- min(p_boot_left, p_boot_right)
cat("One-tailed bootstrap probability value for the null hypothesis of symmetry: p =", round(p_boot, 3),
"\n")
```

```
if (p_boot < alpha) {cat(sprintf("The null hypothesis of normality is rejected at a significance level of %.2f
using bootstrap probability.", alpha), "\n")
} else {cat(sprintf("The null hypothesis of normality is maintained at a significance level of %.2f using
bootstrap probability.", alpha), "\n")}
```

5.2.3. The output of the script applied to the vector of scores x

The 60 sample data points, shown in blue in the script, were simulated from a Cauchy distribution with parameters: x_0 = 0.05 (location parameter corresponding to the median) and γ = 0.38 (scale parameter corresponding to the semi-interquartile range) [73]. To give the data a concrete context within the field of social sciences, they are considered to correspond to stock price return data in volatile markets. For this type of data, the Cauchy distribution probability model has been found to exhibit good predictive ability [46-47]. The data were rounded to three decimal places. Next, the script for random sample generation is presented, which is stable due to having a fixed seed [74].

```
set.seed(123)
x <- rcauchy(60, location = 0.05, scale = 0.38)
print(round(x, 3))
```

Considering a significance level of 5%, the sample of variable X, composed of 60 data, is random by the Wald-Wolfowitz runs test (r = 27, n_0 = 30, n_1 = 30, $z(r)$ = -1.042, two-tailed exact probability value: p = 0.298, ϕ = 0). It deviates from normality by the D'Agostino-Berlanger-D'Agostino ($k^2(x)$ = 17.457, p < 0.001, ϕ = 0.9705), Lilliefors (d = 0.194, p < 0.001, ϕ = 1), Anderson-Darling (AD = 3.4071, p < 0.001, ϕ = 1), and Shapiro-Francia (w = 0.838, p < 0.001, ϕ = 1) tests. In the histogram with the overlaid density and normal curves (Figure 13), a leptokurtic profile is seen, rather symmetrical and with a peak in the center of the distribution. It is thinned in the shoulder zone, but thickened in the tail zone.

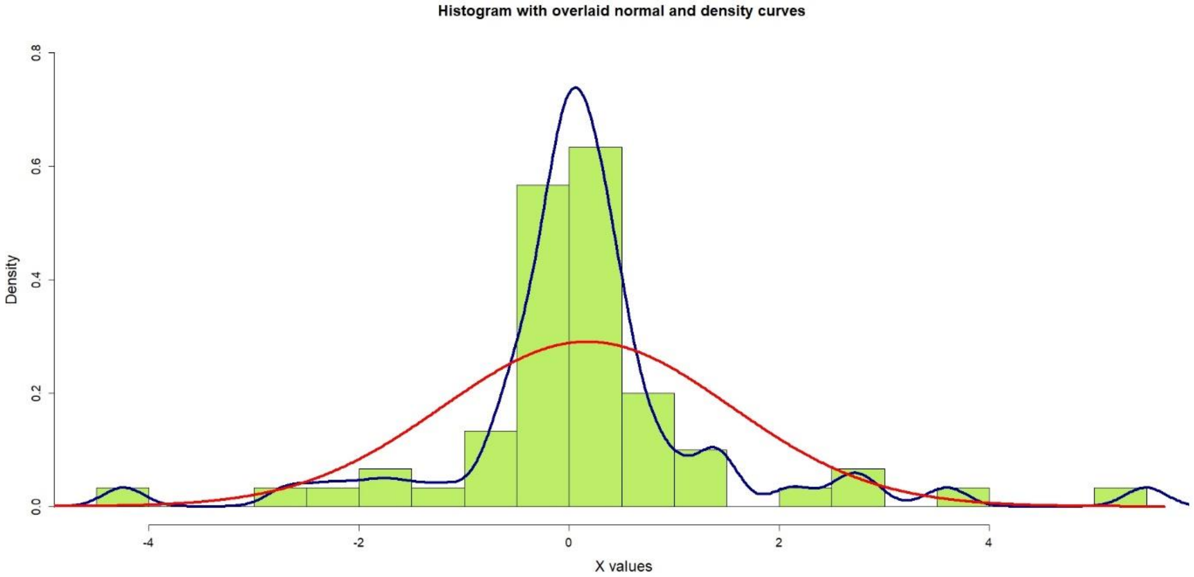

Figure 13. Histogram of X values with overlaid density and normal curves. Source: prepared by the author from the script in sub-section 5.2.2, executed with the R program.

The reversed absolute index of skewness (*AISr*) reaches a value of 0.106, close to 0 (Table 4). When the asymptotic confidence interval is calculated at the 95% level, 0 is within the interval (95% *CI* (-0.267, 0.479); *ase* = 0.190), indicating expected symmetry. However, since the sample of X values does not conform to a normal distribution and is assumed to follow a Cauchy distribution, which lacks finite moments, such an interval is inappropriate and should not be reported. Instead, bootstrap estimation should be used, with a small standard error (*bse* = 0.430), making it reliable.

The bootstrap confidence intervals for the three methods include zero, with the most efficient, having the smallest width, obtained through the percentile method. Its width measures 1.606, compared to

1.664 for the BCa method and 1.685 for the normal method (Table 5). At a 5% significance level, the bootstrap sampling distribution (*b$t*) exhibits symmetry according to the D'Agostino test and mesokurtosis according to the Anscombe-Glynn test, while conforming to normality according to the D'Agostino-Berlanger-D'Agostino test (Table 6). The histogram displays a bell-shaped profile, and the normal quantile-quantile plot shows good alignment of the data at 45 degrees, except for slight deviations at the extreme points (Figure 14). The Shapiro-Francia test confirms normality at a 1% significance level (Table 6), and the skewness and kurtosis values (Table 4) exceed 0.05. Therefore, the BCa confidence interval is deemed the most appropriate (Table 5). Consistent with the bootstrap confidence interval, the null hypothesis of normality is upheld at a 5% significance level using bootstrap probability (Table 5).

Table 4. Point estimation, bootstrap estimation, bootstrap bias, bootstrap standard error, and jackknife acceleration.

Statistic	point_est	boot_est	boot_bias	boot_se	jack_a
AISr	0.106	0.158	0.052	0.430	-0.053
PCS	0.086	0.115	0.029	0.295	-0.055
RISr	1.174	4.470	3.296	71.923	-0.038

[1] Note. Statistic: *AISr* = Reversed Absolute Index of Skewness, *CPS* = Percentile Coefficient of Skewness, and *RISr* = Reversed Relative Index of Skewness, point_est = statistic value or point estimation in the original sample (t_0), boot_est = mean(*b$t*) = bootstrap estimation or mean of the statistic in the 1000 bootstrap samples, boot_bias = boot_est - t_0 = bootstrap = bootstrap bias estimation or difference between the bootstrap estimate and point estimation in the original sample, boot_se = *sd*(*b$t*) = bootstrap standard error or standard deviation of the statistic in the 1000 bootstrap samples, jack_a = jackknife acceleration or measure of the rate of change in the standard deviation ($\sigma[t_0]$) of the estimate ($t_0 = f(x)$) as the data changes. Source: prepared by the author from the script in sub-section 5.2.2, executed with the R program.

Table 5. Bootstrap confidence interval for the three statistics and bootstrap probability for null hypothesis of symmetry.

Statistic of skewness	Bootstrap confidence interval at 95%			Bootstrap p_value
	norm	perc	BCa	
AISr	(-0.789, 0.896)	(-0.629, 0.977)	(-0.742, 0.922)	0.356
PCS	(-0.521, 0.634)	(-0.468, 0.612)	(-0.531, 0.556)	0.356
RISr	(-143.09, 138.84)	(-115.42, 158.42)	(-219.02, 51.65)	0.336

[1] Note. Statistic: *AISr* = Reversed Absolute Index of Skewness, *CPS* = Percentile Coefficient of Skewness, and *RISr* = Reversed Relative Index of Skewness. Confidence intervals using the normal (norm), percentile (perc), and bias-corrected and accelerated percentile (BCa) methods. Calculations based on 1000 extractions with replacement from the original sample. Probability values less than 0.05, indicating the statistical test's significance, are highlighted in bold. Source: prepared by the author from the script in sub-section 5.2.2, executed with R.

Table 6. Skewness and normality in the bootstrap sampling distribution of the three statistics of skewness.

Statistic of	D'Agostino: skewness test			Anscombe-Glynn: kurtosis test			DBD test		Shapiro-Francia Royston	
skewness	$\sqrt{b_1}$	z	p	b_2	z	p	k^2	p	w	p
AISr	0.108	1.400	0.162	2.846	-0.976	0.329	2.912	0.233	0.996	**0.010**
PCS	-0.233	-2.989	**0.003**	2.240	-8.221	**<0.001**	76.522	**<0.001**	0.982	**<0.001**
RISr	1.282	13.135	**<0.001**	44.534	18.365	**<0.001**	509.83	**<0.001**	0.441	**<0.001**

[1] Note. Statistic: *AISr* = Reversed Absolute Index of Skewness, *CPS* = Percentile Coefficient of Skewness, and *RISr* = Reversed Relative Index of Skewness, D'Agostino skewness test: $\sqrt{b_1}$ = measure of skewness based on the third standardized central moment, z = test statistic or standardized value of $\sqrt{b_1}$, *p* = probability value in a two-tailed test. Anscombe-Glynn kurtosis test: b_2 = measure of kurtosis based on the fourth standardized central moment, *z* = test statistic or standardized value of b_2, *p* = probability value in a two-tailed test, D'Agostino-Berlanger-D'Agostino DBD normality test: k^2 = test statistic, *p* = probability value, Shapiro-Francia normality test using Royston's procedure: *w* = test statistic based on the square of the correlation between theoretical and empirical quantiles and *p* = probability value. Probability values less than 0.05, indicating the statistical test's significance,

are highlighted in bold. Source: prepared by the author from the script in sub-section 5.2.2, executed with the R program.

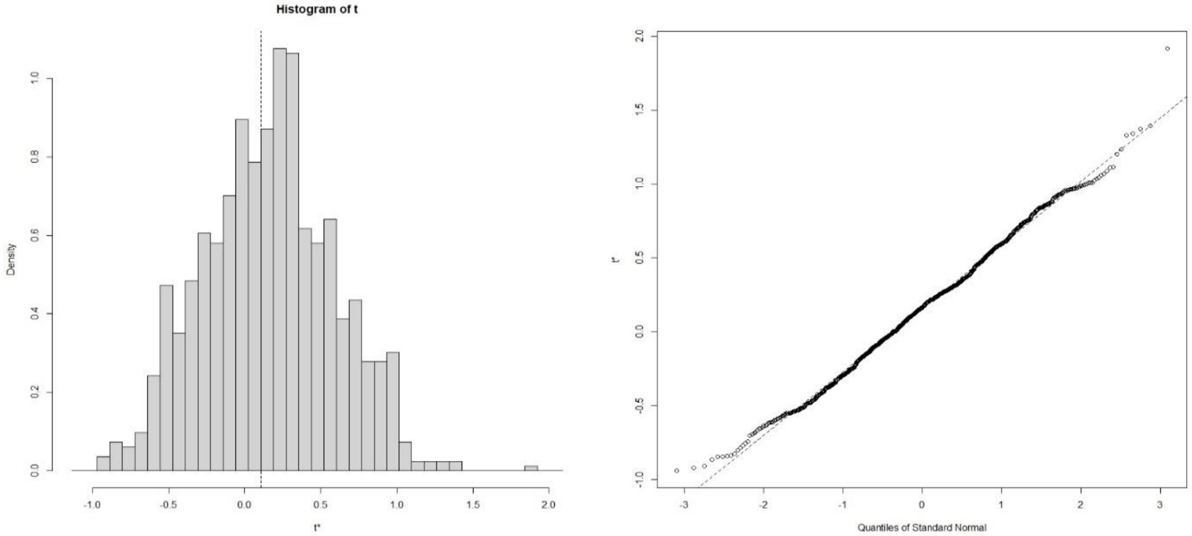

Figure 14. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution for *AISr* derived from the 60-data sample of variable X. Source: prepared by the author from the script in sub-section 5.2.2, executed with the R program.

The percentile coefficient of skewness reached a value of 0.086, close to 0 (Table 4). The asymptotic confidence interval includes 0 (95% *CI*: (-0.217, 0.389), *ase* = 0.155). However, given the non-normal distribution assumed to follow a Cauchy distribution, this interval is not appropriate. All three bootstrap confidence intervals include 0 and are reliable, with small standard errors (*bse* = 0.295) and bias (0.029 < 0.05). The interval with the smallest width is obtained by the percentile method (w = 1.080), followed closely by the BCa method (*w* = 1.087), and finally the normal method (*w* = 1.155) (Table 5). The bootstrap sampling distribution of the statistic shows negative skewness and platykurtosis (Table 6), resulting in a histogram with thickened shoulders and shortened tails, and an elongated S-shape in the normal quantile-quantile plot (Figure 15). Given the bootstrap sampling distribution's deviation from normality due to negative skewness and high kurtosis ($|a|$ = 0.055 > 0.025), the BCa method is considered the most appropriate (Table 5). Consistent with the bootstrap confidence interval, the null hypothesis of normality is upheld at a significance level of 0.05 using bootstrap probability (Table 5).

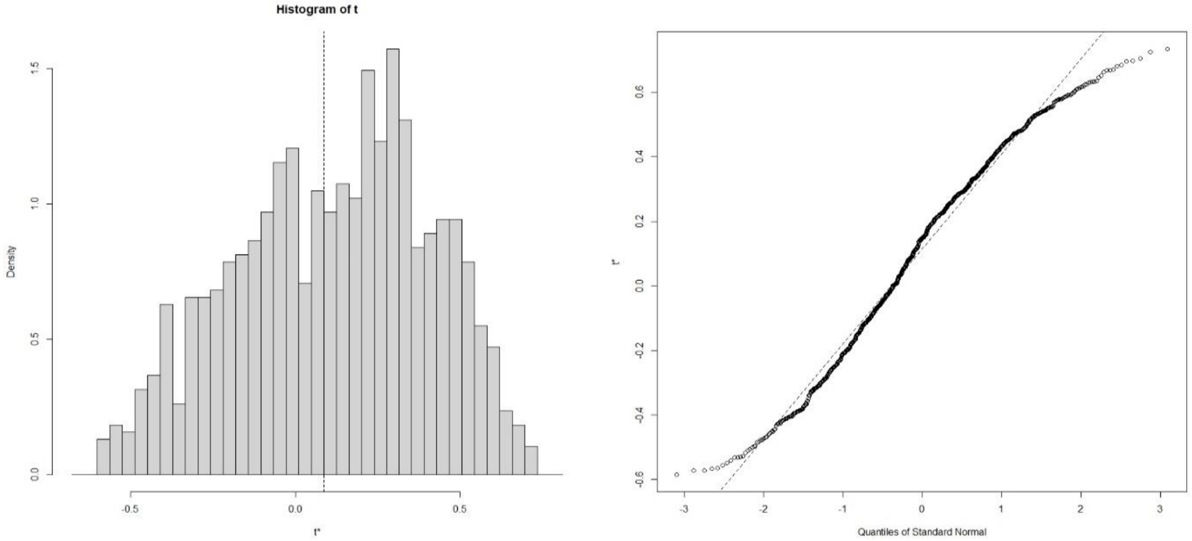

Figure 15. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution for *PCS* derived from the 60-data sample of variable X. Source: prepared by the author from the script in sub-section 5.2.2, executed with the R program.

5.3. Calculation of *RISr* in the sample of 60 data points using R

In the script, the option to compute the bootstrap confidence interval using the BCa method with the 'bcajack' library [72] was included due to potential memory limitations or timeouts when using the 'boot' library. The acceleration value is derived from computations with the 'bcajack' library, but the display of the resulting bootstrap confidence interval using the BCa method is disabled: # print(bcajack_results). To activate it, uncomment boot.ci(b, conf = 0.95, type = c("bca", "perc", "norm")) by adding the symbol # in front and remove it from print(bcajack_results). This package requires significantly fewer computational resources compared to the 'boot' library for computing the bootstrap confidence interval using the BCa method.

```
x <- c(0.532, -0.248, 1.342, -0.096, -0.022, 0.105, -4.243, -0.084, -2.281, 2.821, -0.002, 2.623, -0.559, -1.586,
0.177, -0.074, 0.421, 0.101, 0.683, -0.005, -0.087, -0.499, -0.754, 0.043, -0.664, -0.445, -2.677, -1.197, 0.537,
0.239, 0.006, -0.070, -0.506, -0.234, 0.079, 5.489, -0.310, 0.357, 0.641, 0.388, 0.233, 1.431, 1.417, 0.919, 0.247,
0.227, 0.392, 3.590, 0.470, -0.132, 0.105, 2.120, -0.228, 0.203, -1.910, 0.338, 0.211, -0.322, -0.080, 0.963)

library(bcaboot)
library(boot)
library(moments)
library(nortest)

cat("Point estimate of the additive opposite of Kelley's relative index of skewness", "\n")
RISr <- (quantile(x, 0.10, type = 8) + quantile(x, 0.90, type = 8)) / (2 * quantile(x, 0.50, type = 8)) - 1
cat("Opposite additive Kelley's relative index of skewness: RISr(x) =",round(RISr, 3),"\n")

cat("Bootstrap confidence interval of the additive opposite of Kelley's relative index of skewness", "\n")
# Acceleration (jackknife)
set.seed(123)
bcajack_results <- bcajack(x, B = 1000, function(x, i) {(quantile(x[i], 0.10, type = 8) + quantile(x[i], 0.90,
type = 8)) / (2* quantile(x[i], 0.50, type = 8)) - 1}, alpha = 0.025, verbose = TRUE)
acel <- bcajack_results$stats[1, "a"]
cat("Acceleration (jackknife): a =", round(acel, 6), "\n")
set.seed(123)
b <- boot(data=x, function(x, i) {(quantile(x[i], 0.10, type = 8)+quantile(x[i], 0.90, type = 8))/(2*
quantile(x[i], 0.50, type = 8)) - 1}, R=1000)
boot_est <- mean(b$t)
cat("Bootstrap estimation: m(b$t) =", round(boot_est, 3), "\n")
agostino_boot <- agostino.test(b$t, alternative = "two.sided")
ag_boot <- anscombe.test(b$t, alternative = "two.sided")
print(agostino_boot)
print(ag_boot)
cat("D'Agostino-Berlanger-D'Agostino normality test", "\n")
K2 <- (agostino_boot$statistic["z"]^2 + ag_boot$statistic["z"]^2)
p <- pchisq(K2, df = 2, lower.tail = FALSE)
cat("K^2 =", round(K2, 3), "\n")
cat("p_value =", round(p, 3), "\n")
sf.test(b$t)
b
plot(b)
boot.ci(b, conf = 0.95, type = c("bca", "perc", "norm"))
# print(bcajack_results) # Remove the '#' symbol if the 'boot' library cannot calculate the BCa confidence
interval
```

```
p_boot_izq <- mean(b$t < 0)
p_boot_der <- mean(b$t > 0)
p_boot <- min(p_boot_izq, p_boot_der)
cat("One-tailed bootstrap probability value for the null hypothesis of symmetry: p =", round(p_boot, 3),
"\n")
alpha <- 0.05
if (p_boot < alpha) {cat(sprintf("The null hypothesis of normality is rejected at a significance level of %.2f
using Bootstrap probability.", alpha), "\n")
} else {cat(sprintf("The null hypothesis of normality is maintained at a significance level of %.2f using
bootstrap probability.", alpha), "\n")}
```

Kelley's reversed relative index of skewness (*RISr*) reaches a value of 3.296 (Table 4). The bootstrap sampling distribution of the *RISr* statistic (*b$t*) exhibits positive skewness and strong leptokurtosis, indicating departure from normality (Table 6). The histogram displays a pointed profile with thinned shoulders and elongated tails. The normal quantile-quantile plot flattens in the central area and deviates on both sides, forming a concave curvature to the left and a convex curvature to the right, describing a broad-shouldered shape (Figure 16). Moreover, it shows considerable bias and an acceleration greater than 0.025 (Table 4), making the BCa method the most appropriate for the bootstrap confidence interval, which includes 0 at a 95% level (Table 5). The interval width is notably large (w = 270.67), similar to intervals obtained with the other two methods: 273.84 with the percentile method and 281.93 with the normal method (Table 5), due to a very large bootstrap standard error (*bse* = 71.923) (Table 4), rendering it unreliable. Consistent with the bootstrap confidence interval, the null hypothesis of normality is upheld at a 5% significance level using bootstrap probability (Table 5), although this statistic should ultimately be disregarded due to its lack of reliability.

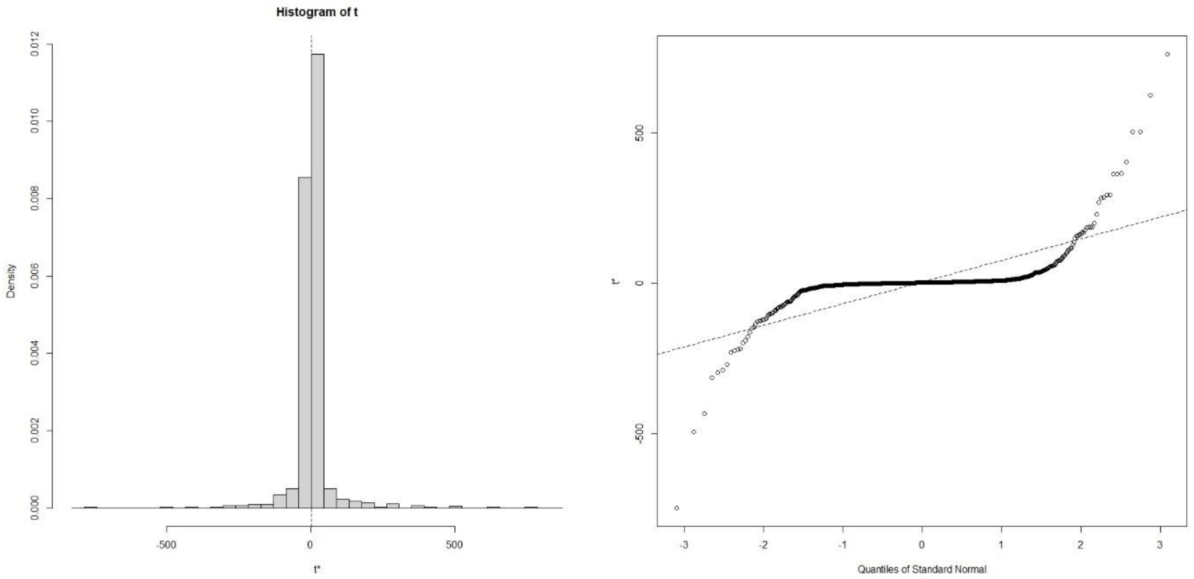

Figure 16. Histogram and normal quantile-quantile plot of the bootstrap sampling distribution for *RISr* derived from the 60-data sample of variable X. Source: prepared by the author from the script in this sub-section 5.3, executed with the R program.

6. Discussion

From the design of this study, source samples of 10,000 data points were generated using inverse transform sampling [37] from three continuous, symmetric distributions with different kurtosis levels (platykurtic, mesokurtic, and leptokurtic). The bootstrap sampling distribution of the absolute index and percentile coefficient of skewness converges to a normal distribution, whereas that of the relative index of skewness deviates due to pronounced leptokurtosis. Therefore, when developing the R script to obtain point and interval estimates for these indices, it includes errors and asymptotic intervals for the absolute index and percentile coefficient. The script allows for checking normality and considering the sample distribution of variable X to choose between asymptotic or bootstrap alternatives. In cases of normality or near-normality (symmetry, mesokurtosis, and a bell-shaped profile) with a large sample size, the asymptotic confidence interval is highly suitable. It's worth noting that asymptotic and bootstrap standard errors for these two statistics are very similar across the three simulated distributions, all of which are symmetric and have finite moments.

Furthermore, the script allows computation of the bootstrap confidence interval for these two measures of skewness using three methods: normal, percentile, and BCa [25, 69]. To determine the most suitable method, the script checks the symmetry, kurtosis, and normality of the bootstrap sampling distribution, along with acceleration, skewness, and error statistics. The normal method is preferred if the distribution is normal [69]. If the distribution is non-normal but exhibits symmetry, low skewness, and low acceleration (< 0.05 and < 0.025, respectively), the percentile method is suitable, especially if it provides a narrower interval than the BCa method. For non-normal distributions with significant skewness and kurtosis, along with substantial bias and acceleration, the BCa method is recommended [72]. If the error is very large—greater than a quarter of the variable's range or the sample size—the interval is questionable and should be approached with caution, or better discarded together with point estimate [23, 25-26].

In the example presented, originating from a distribution without finite moments [46-47], the asymptotic approximation cannot be applied. Moreover, due to the excessive bootstrap standard error, the *RISr* should be disregarded. Given the non-small bias and acceleration, and the less-than-ideal fit to normality of the bootstrap sampling distribution, the BCa method is the optimal choice for obtaining confidence intervals [72]. This method provides intervals with amplitudes very similar to those from the percentile method, which would otherwise be somewhat inadequate [25]. It can be concluded that both indices are reliable, given their small standard errors, and that the sample stems from a symmetric distribution, as evidenced by bootstrap confidence intervals encompassing 0 and bootstrap probability values exceeding the significance level of 0.05 (*AISr* = 0.052, 95% BCa bootstrap *CI* (-0.742, 0.922), two-tailed bootstrap *p*-value = 0.356; *PCS* = 0.029, 95% BCa bootstrap *CI* (-0.531, 0.556), two-tailed bootstrap *p*-value = 0.356). This significance level is appropriate for a sample size of 60, allowing for valid inference due to its random nature. Specifically, this conclusion holds true because the sample originates from a symmetric distribution, such as the Cauchy distribution [46].

Limitations of the study include the exclusion of asymmetric distributions with finite moments, such as gamma, loglogistic, or lognormal distributions [16]. Additionally, we did not vary the size of the source sample to assess convergence to normality (e.g., increments from 30 to 900 by 100, and from 1000 to 10000 by 1000). Instead, the study focused on three symmetric distributions with varying kurtosis and very large source samples, chosen to effectively represent the source distribution [17].

The question arises whether the script is applicable to ordinal variables. If the variable has a large number of ordered categories and assumes normality, a common practice in psychology and other social sciences, it could be feasible [75]. These skewness measures rely on quantiles, which are calculated robustly using R's method 8. However, using a rounding method (such as the 1 to 3 rules in R) for quantile calculation warrants further study [14, 20].

7. Conclusions

The sampling distribution of Kelley's absolute index of skewness converges to a normal distribution. Its additive complement, divided by the semi-percentile rank which corresponds to the coefficient of percentile skewness, serves as its appropriate standardization. The sampling distribution of this coefficient also converges to a normal distribution. It is inappropriate to standardize the absolute index using the median. The R script developed can be practically and didactically useful for point and interval estimation of both the absolute index and its standardized form.

Funding: This research did not receive any external funding.

Acknowledgments: The author expresses gratitude the reviewers and editor for their helpful comments.

Conflicts of Interest: The author declares no conflict of interest.

References

[1] T. L. Kelley, "A new measure of dispersion," Quar. Pub. Amer. Statist. Assoc., vol. 17, no. 134, pp. 743-749, June 1921. https://doi.org/10.1080/15225445.1921.10503833

[2] T. L. Kelley, Statistical Method, The Macmillan Company, New York, 1923, pp. 75-77. https://doi.org/10.1080/15225445.1921.10503833

[3] K. Pearson, "Contributions to the mathematical theory of evolution. I. On the dissection of asymmetrical frequency curves". Phil. Trans. Roy. Soc. London A, vol. 185, pp. 71−110, January 1894. https://doi.org/10.1098/rsta.1894.0003

[4] K. Pearson, "Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material", Phil. Trans. Roy. Soc. London A, vol. 186, pp. 343−414, 1895. https://doi.org/10.1098/rsta.1895.0010

[5] A. L. Bowley, Elements of Statistics, P. S. King and Son, London, 1901.

[6] U. G. Yule, An introduction to the Theory of Statistics. Charles Griffin and Company Limited, London, 1912. https://doi.org/10.1037/13786-000

[7] R. A. Fisher, "The moments of the distribution for normal samples of measures of departure from normality". Proc Roy. Soc. London A, vol. 130, no. 812, pp. 16–28, December 1930. https://doi.org/10.1098/rspa.1930.0185

[8] D. Stout, "A question of statistical inference: E. G. Boring, T. L. Kelley, and the probable error". Am. J. Psychol., vol. 102, no. 4, pp. 549–562, April 1989. https://doi.org/10.2307/1423307

[9] R. Ihaka, R: past and future history, 2022. https://cran.r-project.org/doc/html/interface98-paper/paper.html

[10] D. R. Bickel, "Robust estimators of the mode and skewness of continuous data," Comput. Stat. Data Anal., vol. 39, no. 2, pp. 153−163, April 2002. https://doi.org/10.1016/S0167-9473(01)00057-3

[11] G. Altinay, A simple class of measures of skewness. Munich Personal RePEc Archive, Paper No. 72353, pp. 1−13, September 2016. https://mpra.ub.uni-muenchen.de/72353/

[12] A. Singh, L. Gewali, and J. Khatiwada, "New measures of skewness of a probability distribution," Open J. Stat., vol. 9, no. 3, pp. 601−621, October 2019. http://dx.doi.org/10.4236/ojs.2019.95039

[13] A. Eberl, and B. Klar, "Asymptotic distributions and performance of empirical skewness measures," Comput. Stat. Data Anal., vol. 146, article 106939, June 2020. https://doi.org/10.1016/j.csda.2020.106939

[14] R Development Core Team, Quantile {stats}. R Documentation. Sample Quantiles, 2024. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/quantile.html

[15] K. Stapor, "Descriptive and inferential statistics," in Introduction to Probabilistic and Statistical Methods with Examples in R, vol 176, Intelligent Systems Reference Library, Cham, Switzerland: Springer, 2020, pp. 63–131. https://doi.org/10.1007/978-3-030-45799-0_2

[16] S. C. Gupta, and V. K. Kapoor, Fundamentals of Mathematical Statistics (12 th ed.), New Delhi: Sultan Chand & Sons, 2020.

[17] L. M. Chihara, and T. C. Hesterberg, Mathematical Statistics with Resampling and R. New York: John Wiley & Sons, 2022.

[18] A. Linden, "CENTILE2: Stata module to enhance centile command and provide additional definitions for computing sample quantiles," Statistical Software Components S459262, Boston College Department of Economics, 2023.

[19] D. I.Sukhoplyuev, and A. N. Nazarov, Methods of descriptive statistics in telemetry tasks, in Proceedings of the 2024 Systems of Signals Generating and Processing in the Field of on-Board Communications, Moscow, Russian Federation, (article 10496798), New Orleans, LA: Institute of Electrical and Electronics Engineers (IEEE), 2024. https://doi.org/10.1109/IEEECONF60226.2024.10496798

[20] R. J. Hyndman, and Y. Fan, "Sample quantiles in statistical packages," Am. Stat., vol. 50, no. 4, pp. 361-365, November 1996. https://doi.org/10.2307/2684934

[21] J. W. Tukey, Exploratory Data Analysis, Addison-Wesley, Reading, MA, 1977.

[22] A. Roques, and A. Zhao, "Association rules discovery of deviant events in multivariate time series: an analysis and implementation of the sax-arm algorithm", Image Processing On Line, vol. 12, pp. 604-624, December 2022. https://doi.org/10.5201/ipol.2022.437

[23] P. M. Dixon, "The bootstrap and the jackknife: describing the precision of ecological indices," in S. Scheiner, Ed., Design and Analysis of Ecological Experiments, London: Chapman and Hall/CRC, 2020, pp. 290-318. https://doi.org/10.1201/9781003059813

[24] R. B. D'Agostino, "Transformation to normality of the null distribution of g1," Biometrika, vol. 57, no. 3, pp. 679–681, December 1970. https://doi.org/10.1093/biomet/57.3.679

[25] B. Efron, and B. Narasimhan, "The automatic construction of bootstrap confidence intervals," J. Comput. Graph. Stat., vol. 29, no. 3, pp. 608–619, March 2020. https://doi.org/10.1080/10618600.2020.1714633

[26] G. Rousselet, C. R. Pernet, and R. R. Wilcox, "An introduction to the bootstrap: a versatile method to make inferences by using data-driven simulations," Meta-Psychology, 7, artícle 2019.2058, December 2023 https://doi.org/10.15626/MP.2019.2058

[27] G. Sánchez-Barajas, and A. Gómez-Navarro, Estadística General Aplicada con Excel [General Statistics Applied with Excel]. Universidad Autónoma Nacional de México, Ciudad de México, 2017.

[28] D. V. Hinkley, "On power transformations to symmetry," Biometrika, vol. 62, no. 1, pp. 101−111, April 1975. https://doi.org/10.2307/2334491

[29] R. Chattamvelli, and R. Shanmugam, “Skewness,” in Descriptive Statistics for Scientists and Engineers. Applications in R, Cham, Switzerland: Springer, 2023, pp. 91-110. https://doi.org/10.1007/978-3-031-32330-0_4

[30] S. Luo and D. Villar, The skewness of the price change distribution: A new touchstone for sticky price models. Journal of Money, Credit and Banking, vol. 53, no. 1, pp. 41-72, September 2021. https://doi.org/10.1111/jmcb.12700

[31] P. L. Mamidi, N. D. Arigela, K.N.V.R. Lakshmi, and A Srilakshmi, “Skewness corrected control charts: a new probability model,” Obstetrics and Gynaecology Forum, vol. 34, no. 3s, pp. 775–779, May 2024. Retrieved from https://obstetricsandgynaecologyforum.com/index.php/ogf/article/view/352

[32] F. Najafi, S. Naderpour, M. Moradinazar, M. Khoramdad, A. Vahedian-Azimi, T. Jamialahmadi, and A. Sahebkar, “Percentiles for anthropometric measures in 11–18 years-old students of 73 developing countries,” Diabetes Metab. Syndr., vol. 14, no. 6, pp. 1957-1962, Nov-Dec 2020. https://doi.org/10.1016/j.dsx.2020.10.002

[33] M. Iseringhausen, I. Petrella, and K. Theodoridis, “Aggregate skewness and the business cycle,” Rev. Econ. Stat., 1-37, October 2023. https://doi.org/10.1162/rest_a_01390

[34] K. Harmenberg, The labor-market origins of cyclical skewness, Technical Report, University of Copenhagen, February 2021. Available at https://www.karlharmenberg.com/papers/skewness_harmenberg.pdf

[35] R. A. Groeneveld, and G. Meeden, “Measuring skewness and kurtosis,” J. Roy. Stat. Soc. Ser. D Statistician, vol. 33, no. 4, pp. 391–399, December 1984. https://doi.org/10.2307/2987742

[36] S. Salgado, F. Guvenen, and N. Bloom, “Skewed business cycles,” National Bureau of Economic Research, Working Paper 26565, December 2019. https://doi.org/10.3386/w26565

[37] S. Chowdhury, Monte Carlo Methods Utilizing Mathematica®: Applications in Inverse Transform and Acceptance-Rejection Sampling, Berlin: Springer Nature, 2023. https://doi.org/10.1007/978-3-031-23294-7

[38] R. Kimberk, “Beta distribution of long memory sequences,” arXiv:2404.05736, March 2024. https://doi.org/10.48550/arXiv.2404.05736

[39] F. Grubbs, “Procedures for detecting outlying observations in samples,” Technometrics, vol. 11, no. 1, pp. 1-21, Febrary 1969. https://doi.org/10.1080/00401706.1969.10490657

[40] F. J. Anscombe, and W. J. Glynn, “Distribution of kurtosis statistics b2 for normal samples,” Biometrika, vol. 70, no. 1, pp. 227-234, April 1983. https://doi.org/10.1093/BIOMET/70.1.227

[41] T. W. Anderson, and D. A. Darling, “Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes,” Ann. Math. Stat., vol. 23, no. 2, pp. 193-212, June 1952. http://dx.doi.org/10.1214/aoms/1177729437

[42] S. S. Shapiro, and R. S. Francia, “An approximate analysis of variance test for normality,” J. Am. Stat. Assoc., vol. 67, no. 337, pp. 215–216, April 1972. https://doi.org/10.1080/01621459.1972.10481232

[43] J. P. Royston, “A toolkit for testing for non-normality in complete and censured samples,” J. Roy. Stat. Soc., Ser. D-Statistician, vol. 42, no. 1, pp. 37-43, March 1993. https://doi.org/10.2307/2348109

[44] R. B. D’Agostino, A. Berlanger, and R. B. Jr. D’Agostino, “A suggestion for using powerful and informative tests of normality,” Am. Stat., vol. 44, no. 4, pp. 316-321, November 1990. https://doi.org/10.2307/2684359

[45] P. Kvam, B. Vidakovic, and S. J. Kim, Nonparametric Statistics with Applications to Science and Engineering with R, 2nd ed., Hoboken, NJ: John Wiley & Sons, 2022. https://doi.org/10.1002/9781119268178

[46] A. Blöchlinger, Gauss versus Cauchy: a comparative study on risk, in T. Hüttche, Ed., Finance in Crises. Contributions to Finance and Accounting, Cham, Switzerland: Springer, 2023, pp. 177–198. https://doi.org/10.1007/978-3-031-48071-3_12

[47] S. Nadarajah, and T. Hitchen, “Estimation of models for stock returns,” Computational Economics, March 2024. https://doi.org/10.1007/s10614-024-10580-x

[48] B. Coker, C. Rudin, and G. King, “A theory of statistical inference for ensuring the robustness of scientific results,” Management Science, vol. 67, no. 10, pp. 6174-6197, October 2021. https://doi.org/10.1287/mnsc.2020.3818

[49] F. Caeiro, and A. Mateus, Randtests: Testing Randomness in R, 2024. https://cran.r-project.org/web/packages/randtests/index.html

[50] S. Kwak, “Are only p-values less than 0.05 significant? A p-value greater than 0.05 is also significant!,” J. Lipid. Atheroscler., vol. 12, no. 2, pp. 89–95, May 2023. https://doi.org/10.12997/jla.2023.12.2.89

[51] G. Di Leo, and F. Sardanelli, “Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach,” Eur. Radiol. Exp., vol. 4, article 18, March 2020. https://doi.org/10.1186/s41747-020-0145-y

[52] J. H. Kim, and I. Choi, “Choosing the Level of Significance: A Decision-theoretic Approach,” Abacus, vol. 57, no. 1, pp. 27-71, March 2021. https://doi.org/10.1111/abac.12172

[53] K. J. Nicholson, M. Sherman, S. N. Divi, D. R. Bowles, and A. R. Vaccaro, “The Role of Family-wise Error Rate in Determining Statistical Significance,” Clin. Spine Surg., vol. 35, no. 5, pp. 222-223, June 2022. https://doi.org/10.1097/BSD.0000000000001287

[54] C. Avram, and M. Mărușteri, “Normality assessment, few paradigms and use cases,” Rev. Romana Med. Lab., vol. 30, no. 3, pp. 251-260, July 2022. https://doi.org/10.2478/rrlm-2022-0030

[55] L. Komsta, and F. Novomestky, Moments: moments, cumulants, skewness, kurtosis and related tests, 2022. https://cran.r-project.org/web/packages/moments/index.html

[56] J. Gross, and U. Ligges, Package ‘nortest’, 2022. https://cran.r-project.org/web/packages/nortest/nortest.pdf

[57] H. W. Lilliefors, “On the Kolmogorov-Smirnov test for normality with mean and variance unknow,” J. Am. Stat. Assoc., vol. 62, no. 318, pp. 399-402, August 1967. https://doi.org/10.2307/2283970

[58] S. Demir, “Comparison of normality tests in terms of sample sizes under different skewness and Kurtosis coefficients,” Int. J. Assess. Tool. Educ., vol. 9, no. 2, pp. 397-409, May 2022. https://doi.org/10.21449/ijate.1101295

[59] N. Khatun, “Applications of Normality Test in Statistical Analysis,” Open J. Stat., vol. 11, no. 1, pp. 113-122, February 2021. https://doi.org/10.4236/ojs.2021.111006

[60] I. Mala, V. Sladek, and D. Bılkova, “Power comparisons of normality tests based on l-moments and classical tests,” Math. Stat., vol. 9, no. 6, pp. 994-1003, November 2021. https://doi.org/10.13189/ms.2021.090615

[61] D. K. Wijekularathna, A. B Manage, and S. M. Scariano, “Power analysis of several normality tests: A Monte Carlo simulation study,” Commun. Stat-Simul C Journal, vol. 51, no. 3, pp. 757-773, September 2020. https://doi.org/10.1080/03610918.2019.1658780

[62] R. Sahann, T. Müller, and J. Schmidt, “Histogram binning revisited with a focus on human perception,” in Proceeding of the 2021 IEEE Visualization Conference (VIS), New Orleans, LA: Institute of Electrical and Electronics Engineers (IEEE), 2021, pp. 66-70. https://doi.org/10.1109/VIS49827.2021.9623301

[63] V. A. Epanechnikov, “Nonparametric estimation of a multidimensional probability density,” Theory of Probability and its Applications, vol. 14, no. 1, pp. 156–161, January 1969. https://doi.org/10.1137/1114019

[64] N. Fadillah, P. A. Dariah, A. Anggraeni, N. Cahyani, and L. Handayani, “Comparison of Gaussian and Epancehnikov Kernels,” Tadulako Social Science and Humaniora Journal, vol. 3, no. 1, pp. 13-22, September 2022. https://doi.org/10.22487/sochum.v3i1.15745

[65] S. J. Sheather, and M. C. Jones, “A reliable data-based bandwidth selection method for kernel density estimation,” J. Roy. Stat. Soc., Ser. B Stat. Methodol., vol. 53, no. 3, pp. 683–690, January 1991. https://doi.org/10.1111/j.2517-6161.1991.tb01857.x

[66] W. N. Venables, and B. D. Ripley, Modern Applied Statistics with S, 4th ed., Cham, Switzerland: Springer, 2002. https://doi.org/10.1007/978-0-387-21706-2

[67] E. C. Ogwu, and H. I. A. Ojarikre, “Comparative study of the rule of thumb, unbiased cross validation and the Sheather Jones-direct plug-in approaches of kernel density estimation using real life data,” Int. J. Innov. Res. Sci. Eng. Technol., vol. 11, no. 3, pp. 1800-1809, March 2023.

[68] R Development Core Team. The R Manuals, 2024. https://cran.r-project.org/manuals.html

[69] A. Canty, B. Ripley, and A. R. Brazzale, Package ‘boot’, 2024. https://cran.r-project.org/web/packages/boot/boot.pdf

[70] G. Cavaliere, and I. Georgiev, “Inference under random limit bootstrap measures,” Econometrica, vol. 88, no. 6, pp. 2547-2574, November 2020. https://doi.org/10.3982/ECTA16557

[71] D. Lakens, “Sample size justification,” Collabra: Psychology, vol. 8, no. 1, article 33267, January 2022. https://doi.org/10.1525/collabra.33267

[72] B. Efron, and B. Narasimhan, Package ‘bcaboot’. Bias corrected bootstrap confidence intervals, 2022. https://cran.r-project.org/web/packages/bcaboot/bcaboot.pdf

[73] A. Pekgör, “A novel goodness-of-fit test for Cauchy distribution,” J. Math., vol. 2023, no. 1, article 9200213, March 2023. https://doi.org/10.1155/2023/9200213

[74] B. Lawson, L. Leemis, and V. Kudlay, Set.seed: Seeding random variate generators, 2023. https://www.rdocumentation.org/packages/simEd/versions/2.0.1/topics/set.seed

[75] T. L. Kelley, Fundamentals of Statistics, Cambridge, MA: Harvard University Press, 1947.