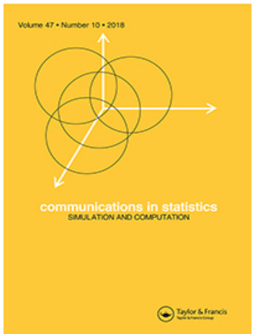

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Simulation and Computation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Confidence intervals for means and variances of nonnormal distributions

José Dias Curto

To cite this article: José Dias Curto (2021): Confidence intervals for means and variances of nonnormal distributions, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2021.1963448

To link to this article: https://doi.org/10.1080/03610918.2021.1963448

Published online: 16 Aug 2021.

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2021.1963448

Taylor & Francis
Taylor & Francis Group

Confidence intervals for means and variances of nonnormal distributions

José Dias Curto[a,b]

[a]Instituto Universitário de Lisboa (ISCTE-IUL), BRU-UNIDE, Lisboa, Portugal; [b]Department of Quantitative Methods, Complexo INDEG/ISCTE, Lisboa, Portugal

ABSTRACT

In this article, we propose new confidence intervals for the population mean and variance, the ratio of two populations variance, and the difference in the arithmetic averages of two populations with nonnormal distribution. Theoretical and practical aspects of the suggested techniques are presented, as well as their comparison with existing methods based on the estimated coverage probability. The suggested confidence intervals give consistent and best coverage in comparison with other methods. In addition, application of presented methods to a data set in domain of auditing and accounting is described and analyzed. The empirical results confirm the Monte Carlo simulation studies, highlighting the superiority of the now proposed methods.

ARTICLE HISTORY
Received 29 September 2020
Accepted 28 July 2021

KEYWORDS
Coverage probability; geometric Average; Monte Carlo simulation

MATHEMATICAL SUBJECT CLASSIFICATION
62E20; 65C05; 62P25

1. Introduction

A confidence interval for population parameters gives the bounds where it is expected that parameters lie with a certain confidence level. In this paper, following the works of Tan and Gleser (1993), Chen (1995), Chen and Chen (1999), Cojbasic and Tomovic (2007) and Cahoy (2010), and based on the results of Bonett (2006), Shoemaker (2003) and Feng et al. (2013), we derive new confidence intervals, with coverage probability closer to the confidence level, for the population mean and variance, the ratio of two population variances, and the difference in arithmetic averages of two populations with nonnormal distribution, specially in the case of leptokurtic distributions.

First, we propose a new estimator for the fourth standardized moment about the population central tendency by replacing the arithmetic mean and the trimmed mean proposed by Bonett (2006) with the median and we compare the bias and the coverage probabilities of the three resulting estimators. We show in Sec. 2 that all the three estimators have negative bias in leptokurtic distributions underestimating the true value of kurtosis. The Pearson estimator, the one based on arithmetic mean, has the largest negative bias. From the two estimators based on the trimmed mean and median, the one based on median has the smallest negative bias. Estimated coverage probability of the resulting confidence intervals also confirms its superiority. Thus, it seems the most appropriate to estimate the level of kurtosis of leptokurtic distributions, and this is the first contribution of the paper.

Second, still in Sec. 2, following the methodology of Bonett (2006) and considering the distribution of the sample variance (Mood, Graybill, and Boes 2007) and its logarithm (Shoemaker 2003), we derive a confidence interval for the population variance using the median to estimate the fourth central moment. From the Monte Carlo simulation study, and in case of leptokurtic

CONTACT José Dias Curto dias.curto@iscte-iul.pt Instituto Universitário de Lisboa (ISCTE-IUL), BRU-UNIDE, Lisboa 1600-189, Portugal.

distributions, the estimator based on the median results in confidence intervals for the population variance with estimated coverage probability closer to the used confidence level. Thus, providing a more accurate confidence interval for the population variance in case of leptokurtic distributions is the second contribution of the paper.

Third, we generalize in Sec. 3 the results of Bonett (2006) for one single population variance to the ratio of two population variances and we compare the resulting confidence intervals with the ones based on $F-$ Snedecor distribution with the adjustment in degrees of freedom suggested by Shoemaker (2003). The results show that all the confidence intervals are conservative in platykurtic distributions. For moderate leptokurtic distributions the coverage probability of confidence intervals resulting from the adjustment is closer to the confidence level. In the extreme leptokurtic distribution (χ_1^2) only the coverage probability of the confidence intervals resulting from the new method (based on the median to estimate the fourth central moment) is close to the confidence level. Thus, this more accurate method in case of extreme leptokurtic distributions is the third contribution of the paper.

Many standard statistical and econometrical analyses, such as regression or the analysis of variance, have key assumptions implicit, namely that data is normally distributed (or at least symmetrically distributed), with constant variance. If the evidence indicates that the data assumptions cannot be satisfied, using parametric statistical tests on such data may give a misleading result and two courses of action are available. The first is to carry out a different statistical technique which does not require the violated assumptions, such as non-parametric tests (in case of ANOVA, for example). The second is to transform the data expecting that the "new" data meets the assumptions of the analysis. Once this is accomplished, we can carry out the analysis on the transformed variable. Where possible, data transformation is generally the easier of these two ways (see Box and Cox (1964) and Atkinson (1986)). For right-skewed data, the log transformation is, arguably, the most popular among the different types of transformations. However, several authors including Feng et al. (2013), have found many misuses and misinterpretations of analysis based on log-transformed data. For example, a common practice in statistics is to take the log transformation and construct confidence intervals on the basis of the transformed data. However, when computed based on log-transformed data, the confidence interval is for the geometric mean and not for the arithmetic mean of the original data. Thus, the fourth contribution of this paper is to propose a confidence interval for the population arithmetic mean based on the confidence interval for the population geometric mean resulting from the log-transformed data. The coverage probability of the new confidence interval is very close to the nominal confidence level no matter the sample size (see Sec. 4 for details).

The confidence intervals for the difference of means of two populations appear in Sec. 5. We extend the results of Bonett (2006) and Johnson (1978) and we derive a confidence interval for the ratio of two populations arithmetic averages based on the confidence interval for the ratio of two geometric averages. In terms of coverage probability the intervals give approximately the same results but the lower length is achieved by the new proposed method, providing more efficient estimates. This is the fifth contribution of the paper.

In Sec. 6 we analyze two populations of payables and receivables of a Portuguese company for the year 2019. The empirical results confirm the Monte Carlo simulation studies, highlighting the superiority of the now proposed methods. Finally, in Sec. 7, we present our concluding remarks.

2. Confidence interval for the variance

Let $X_1, X_2, \ldots, X_n$ be a random sample. If $X_i \sim \mathcal{N}(\mu, \sigma^2)$, for all i, an exact $100(1-\alpha)\%$ confidence interval for σ^2 is:

$$\frac{(n-1)\hat{\sigma}^2}{q_2} < \sigma^2 < \frac{(n-1)\hat{\sigma}^2}{q_1}, \tag{1}$$

where $q_1 = \chi^2_{\alpha/2;n-1}, q_2 = \chi^2_{1-\alpha/2;n-1}$ and $\chi^2_{q;df}$ represent the quantiles of the chi-squared distribution with df degrees of freedom. Taking the square root of the endpoints of (1) gives a confidence interval for σ.

As the confidence interval is very sensitive to minor violations of the normality assumption, next we propose and discuss alternatives to the exact case. Let $X_i (i = 1, 2, \dots, n)$ be continuous, independent and identically distributed random variables with finite mean, variance and fourth moment. According to Mood, Graybill, and Boes (2007), Shoemaker (2003) and Bonett (2006), the variance of $\ln(\hat{\sigma}^2)$ is given by (with a small-sample adjustment):

$$V[\ln(\hat{\sigma}^2)] \cong \frac{1}{n-1}\left[\gamma_4 - \frac{(n-3)}{n}\right], \tag{2}$$

where $\gamma_4 = \frac{\mu_4}{\sigma^4}$, μ_4 is the fourth moment about the population mean and σ is the standard deviation. In practice, γ_4 and σ are unknown, and their usual estimators are:

$$\hat{\gamma}_4(1) = \frac{n\sum_{i=1}^n (X_i - \hat{\mu})^4}{\left[\sum_{i=1}^n (X_i - \hat{\mu})^2\right]^2}, \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \hat{\mu})^2}{n-1} \quad \text{and} \quad \hat{\mu} = \frac{\sum_{i=1}^n X_i}{n}. \tag{3}$$

As $\hat{\gamma}_4(1)$ tends to have large negative bias in leptokurtic distributions, we also consider the estimator $\hat{\gamma}_4(2)$ proposed by Bonett (2006), which is asymptotically equivalent to the Pearson's estimator:

$$\hat{\gamma}_4(2) = \frac{n\sum_{i=1}^n (X_i - \hat{\mu}_m)^4}{\left[\sum_{i=1}^n (X_i - \hat{\mu})^2\right]^2}, \quad \hat{\gamma}_4(3) = \frac{n\sum_{i=1}^n (X_i - \hat{\mu}_{med})^4}{\left[\sum_{i=1}^n (X_i - \hat{\mu})^2\right]^2}, \tag{4}$$

where μ_m is the trimmed mean with trim-proportion equal to $1/[2(n-4)^{1/2}]$. We propose yet another estimator $\hat{\gamma}_4(3)$ replacing the trimmed mean by the median: $\hat{\mu}_{med}$, the one with the smallest negative bias as we show next.

A simulation study was conducted in order to compare the bias of Pearson estimator with the one of the alternative estimators of kurtosis in Eq. (4). We simulate 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 from various theoretical distributions: standard normal: N(0, 1), uniform: U(0, 1), beta: B(3, 3) and B(1, 10), logistic: Log, Laplace: Lap, Student's *t* with 5 degrees of freedom: t(5), gamma: G(1, 6), exponential: Exp, and chi-squared with 1 degree of freedom: χ^2_1. The simulation routines have been programmed in R and they are available if requested.

As one can see in Table 1, the bias of the three estimators is negative (with just a few exceptions) in leptokurtic distributions, confirming the results of Bonett (2006). Thus, on average, the three estimators understate the true value of kurtosis. The Pearson estimator, the most popular, has the largest negative bias. From the two estimators based on trimmed mean and median: $\hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$, respectively, the one based on median has, on average, the smallest negative bias. Thus, it seems the most appropriate to estimate the kurtosis of leptokurtic distributions.

To compare the three estimators[1] for the coefficient of kurtosis: $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$, we also carry out the following simulation by increasing the sample size to 200 and 500. First, we generate 10,000 samples with different size (50, 100, 200 and 500) from the statistical distributions used before. Then we bootstrap, by re-sampling 1,000 times with replacement, each of the 10,000 samples generated before. See, for example, González-Manteiga, Cao, and Marron (1996),

[1]This research topic was suggested by one of the referees.

Table 1. Bias of $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$.

Dist.	n	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$	Dist.	n	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
N(0, 1)	10	−0.5414	−0.3323	0.0532	Lap	10	−2.9471	−2.5375	−2.1471
	20	−0.2826	−0.2131	0.0537		20	−2.0704	−1.8691	−1.5654
	30	−0.1946	−0.1593	0.0355		30	−1.6107	−1.4817	−1.2407
	40	−0.1468	−0.1248	0.0295		40	−1.3478	−1.2571	−1.0575
	50	−0.1170	−0.1017	0.0254		50	−1.1498	−1.0801	−0.9057
	100	−0.0576	−0.0494	0.0152		100	−0.6690	−0.6250	−0.5349
U(0, 1)	10	0.1938	0.3356	0.8891	t(5)	10	−6.1727	−5.8250	−5.4489
	20	0.1210	0.1511	0.5381		20	−5.4053	−5.2218	−4.9383
	30	0.0830	0.0956	0.3797		30	−4.9348	−4.8048	−4.5821
	40	0.0645	0.0713	0.2928		40	−4.6104	−4.5089	−4.3240
	50	0.0518	0.0560	0.2404		50	−4.3788	−4.2959	−4.1349
	100	0.0256	0.0277	0.1238		100	−3.5807	−3.5189	−3.4352
B(3, 3)	10	−0.0885	0.0671	0.4831	G(1, 6)	10	−1.4246	−1.1076	−0.6119
	20	−0.0092	0.0292	0.3099		20	−0.9566	−0.7787	−0.2813
	30	−0.0006	0.0157	0.2203		30	−0.7209	−0.5881	−0.0883
	40	0.0047	0.0137	0.1738		40	−0.5936	−0.4867	0.0143
	50	0.0051	0.0108	0.1416		50	−0.5070	−0.4171	0.0843
	100	0.0036	0.0063	0.0729		100	−0.2738	−0.1804	0.2970
Log	10	−1.5197	−1.2355	−0.8567	Exp	10	−5.8271	−4.9979	−3.9308
	20	−0.9984	−0.8761	−0.6058		20	−4.5549	−3.9186	−2.3262
	30	−0.7529	−0.6788	−0.4720		30	−3.7884	−3.2620	−1.3966
	40	−0.6036	−0.5520	−0.3850		40	−3.2471	−2.7927	−0.7405
	50	−0.5083	−0.4699	−0.3286		50	−2.8617	−2.4615	−0.2823
	100	−0.2899	−0.2672	−0.1965		100	−1.8276	−1.3803	0.9331
B(1, 10)	10	−4.4940	−3.8052	−2.8342	χ^2_1	10	−11.2528	−9.9512	−8.2984
	20	−3.5958	−3.1075	−1.7294		20	−9.2556	−8.2219	−5.6224
	30	−3.1297	−2.7432	−1.1649		30	−7.9967	−7.1329	−4.0416
	40	−2.8403	−2.5179	−0.8180		40	−7.1007	−6.3526	−2.9455
	50	−2.6508	−2.3732	−0.5845		50	−6.3828	−5.7162	−2.0835
	100	−2.2092	−1.9086	−0.0531		100	−4.3967	−3.6377	0.2633

We simulate 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 from various theoretical distributions: standard normal: N(0, 1), uniform: U(0, 1), beta: B(3, 3) and B(1, 10), logistic: Log, Laplace: Lap, Student's t with 5 degrees of freedom: t(5), gamma: G(1, 6), exponential: Exp, and chi-squared with 1 degree of freedom: χ^2_1. $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ are the estimators for the standardized fourth central moment; see Eqs. (3) and (4). The bias is computed as $\hat{\gamma}_4(j) - \gamma_4(j)$, where $\gamma_4(j)$ is the true value of kurtosis. For example, in case of the exponential distribution the kurtosis is 6. The average value is computed based on the 100,000 Monte Carlo samples.

for bootstrap details. The estimate for the coverage probability of the 95% confidence intervals, resulting from percentiles 2.5 and 97.5, is computed based on the proportion of the 10,000 confidence intervals including the true value of the coefficient of kurtosis. The results are shown in Table 2.

If the estimators are consistent, we expect that the bootstrap distributions should collapse around the true value of the kurtosis for the various distributions under analysis.

The performance of $\hat{\gamma}_4(1)$ is getting worse as the distribution moves from N(0, 1), U(0, 1) and B(3, 3) to the leptokurtic distributions. This observation corroborates the results of Kim and White (2004), where, in case of the student's t with 5 degrees of freedom, the center of the boxplot resulting from the Monte Carlo simulation study is still far away from the true value of kurtosis (6, for 5 degrees of freedom) even for $n = 5000$.

Confidence intervals resulting from $\hat{\gamma}_4(3)$ are still liberal in case of leptokurtic distributions, but the difference of the estimated coverage probability to the nominal 95% confidence level is substantially lower, for most of the leptokurtic distributions, when compared to the two other estimators. Thus, the lack of consistency is more evident for $\hat{\gamma}_4(1)$ and $\hat{\gamma}_4(2)$ even when the sample size is $n = 500$, pointing to the best performance of the estimator based on the median: $\hat{\gamma}_4(3)$, in case of leptokurtic distributions.

As we referred before, the main purpose of this section is to propose a new confidence interval for the population variance in case of nonnormal distributions. The exact distribution of the sample variance $\hat{\sigma}^2$ is skewed to the right. Given the desirable properties of $\ln(\hat{\sigma}^2)$, much of the

Table 2. Estimated 95% probabilities of C. I. for the coefficient of kurtosis based on $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$.

Dist.	n	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$	Dist.	n	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
N(0, 1)	50	0.887	0.914	0.915	Lap	50	0.550	0.596	0.656
	100	0.873	0.889	0.921		100	0.597	0.628	0.663
	200	0.889	0.896	0.931		200	0.670	0.685	0.709
	500	0.910	0.912	0.938		500	0.739	0.747	0.759
U(0, 1)	50	0.972	0.968	0.910	t(5)	50	0.326	0.369	0.443
	100	0.964	0.962	0.915		100	0.361	0.387	0.424
	200	0.953	0.954	0.920		200	0.415	0.429	0.452
	500	0.949	0.950	0.931		500	0.502	0.511	0.520
B(3, 3)	50	0.978	0.985	0.979	G(1, 6)	50	0.652	0.732	0.893
	100	0.959	0.963	0.970		100	0.690	0.756	0.894
	200	0.956	0.959	0.969		200	0.740	0.797	0.916
	500	0.956	0.956	0.960		500	0.791	0.836	0.928
Log	50	0.648	0.694	0.820	Exp	50	0.468	0.589	0.784
	100	0.677	0.702	0.773		100	0.530	0.646	0.854
	200	0.725	0.739	0.772		200	0.585	0.694	0.902
	500	0.789	0.795	0.815		500	0.667	0.753	0.892
B(1, 10)	50	0.435	0.582	0.822	χ^2_1	50	0.400	0.518	0.686
	100	0.432	0.574	0.849		100	0.469	0.579	0.774
	200	0.428	0.557	0.884		200	0.517	0.610	0.850
	500	0.437	0.612	0.922		500	0.614	0.690	0.905

We generate 10,000 samples of different sizes: 50, 100, 200 and 100 from various theoretical distributions: standard normal: N(0, 1), uniform: U(0, 1), beta: B(3, 3) and B(1, 10), logistic: Log, Laplace: Lap, Student's *t* with 5 degrees of freedom: t(5), gamma: G(1, 6), exponential: Exp, and chi-squared with 1 degree of freedom: χ^2_1. Then we bootstrap, by re-sampling 1,000 times with replacement, each of the 10,000 samples generated before. $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ are the estimators for the standardized fourth central moment; see Eqs. (3) and (4). The estimated coverage probability is computed based on the proportion of the 10,000 confidence intervals including the true value of the kurtosis.

asymmetry can be removed and hence the normal approximation improved. Thus, large-sample confidence intervals for σ^2 may be obtained from a reverse-transformed confidence interval for $\ln(\sigma^2)$:

$$\exp\left[\ln(\hat{\sigma}^2) \pm z_{1-\alpha/2} se(1)\right] \quad \text{and} \quad \exp\left[\ln(c\hat{\sigma}^2) \pm z_{1-\alpha/2} se(2)\right], \tag{5}$$

where $z_{1-\alpha/2}$ is the quantile of the standardized normal distribution, se is the standard error of $\ln(\hat{\sigma}^2)$:

$$se(1) = \left\{\frac{1}{n-1}\left[\hat{\gamma}_4 - \frac{(n-3)}{n}\right]\right\}^{1/2} \quad \text{and} \quad se(2) = c\left\{\frac{1}{n-1}\left[\hat{\gamma}_4 - \frac{(n-3)}{n}\right]\right\}^{1/2}, \tag{6}$$

respectively, where $c = n/(n - z_{1-\alpha/2})$ is a small-sample adjustment that helps equalize the tails probabilities (Bonett 2006). Taking the square root of the limits of the intervals in (5) gives a confidence interval for σ.

Next we compare the coverage probability of the confidence interval resulting from Eq. (1), that we identify by "Normal" in the Tables 3 and 4, with the coverage probability of the three confidence intervals obtained from Eq. (5), considering the alternative standard errors for $\ln(\hat{\sigma}^2)$ in Eq. (6), and using the three different estimators for the standardized fourth central moment γ_4 in Eqs. (3) and (4), that we represent by $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$, respectively. Estimates of coverage probabilities of (1) and (5) were obtained using 100,000 Monte Carlo random samples of different sizes from the statistical distributions used in the simulation before. For the confidence intervals in (5), the standard error of $\ln(\hat{\sigma}^2)$ is estimated according to the equations in (6) and considering the three different estimators for γ_4. The simulation results are presented in Table 3 (using $se(1)$) and Table 4 (using $se(2)$). The results show, as it was expected, that (1) and (5) have coverage probability close to $\lambda = 1 - \alpha = 95\%$, the confidence level, when sampling from a normal distribution.

Table 3. Estimated 95% probabilities of C. I. in (5) considering standard error *se*(1).

Dist.	n	Normal	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$	Dist.	n	Exact	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
N(0, 1)	10	0.950	0.898	0.908	0.928	Lap	10	0.837	0.809	0.828	0.853
	20	0.951	0.919	0.922	0.936		20	0.817	0.855	0.861	0.877
	30	0.951	0.929	0.930	0.942		30	0.808	0.876	0.879	0.890
	40	0.949	0.933	0.934	0.943		40	0.807	0.891	0.893	0.901
	50	0.950	0.936	0.937	0.944		50	0.799	0.898	0.899	0.906
	100	0.950	0.945	0.945	0.948		100	0.794	0.920	0.921	0.923
U(0, 1)	10	0.992	0.949	0.954	0.970	t(5)	10	0.873	0.821	0.838	0.865
	20	0.996	0.959	0.960	0.977		20	0.842	0.849	0.856	0.874
	30	0.997	0.959	0.960	0.975		30	0.828	0.864	0.868	0.880
	40	0.997	0.958	0.959	0.973		40	0.814	0.875	0.877	0.887
	50	0.997	0.958	0.958	0.971		50	0.806	0.883	0.884	0.892
	100	0.998	0.956	0.956	0.965		100	0.789	0.903	0.904	0.907
B(3, 3)	10	0.978	0.920	0.927	0.947	G(1, 6)	10	0.916	0.862	0.877	0.905
	20	0.981	0.936	0.938	0.953		20	0.906	0.885	0.891	0.914
	30	0.982	0.941	0.942	0.954		30	0.901	0.895	0.899	0.920
	40	0.982	0.943	0.943	0.954		40	0.899	0.905	0.909	0.927
	50	0.982	0.945	0.945	0.954		50	0.898	0.911	0.914	0.932
	100	0.983	0.948	0.948	0.953		100	0.896	0.927	0.930	0.946
Log	10	0.908	0.861	0.874	0.897	Exp	10	0.765	0.719	0.760	0.812
	20	0.896	0.889	0.893	0.909		20	0.731	0.780	0.804	0.862
	30	0.892	0.902	0.905	0.917		30	0.717	0.815	0.831	0.889
	40	0.890	0.911	0.912	0.921		40	0.709	0.835	0.848	0.904
	50	0.888	0.916	0.917	0.924		50	0.703	0.849	0.860	0.914
	100	0.884	0.932	0.933	0.936		100	0.691	0.885	0.895	0.938
B(1, 10)	10	0.829	0.778	0.807	0.854	χ^2_1	10	0.641	0.650	0.711	0.783
	20	0.813	0.832	0.851	0.897		20	0.601	0.733	0.767	0.847
	30	0.806	0.859	0.872	0.918		30	0.587	0.775	0.798	0.877
	40	0.804	0.875	0.886	0.932		40	0.580	0.803	0.820	0.893
	50	0.801	0.886	0.895	0.941		50	0.575	0.820	0.834	0.906
	100	0.798	0.915	0.923	0.960		100	0.562	0.863	0.876	0.932

The coverage probability is computed based on 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 from various theoretical distributions: standard normal: N(0, 1), uniform: U(0, 1), beta: B(3, 3) and B(1, 10), logistic: Log, Laplace: Lap, Student's *t* with 5 degrees of freedom: t(5), gamma: G(1, 6), exponential: Exp, and chi-squared with 1 degree of freedom: χ^2_1. $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ are the estimators for the standardized fourth central moment; see Eqs. (3) and (4). The confidence intervals, except "Normal", are computed based on *se*(1), the standard error of $\ln(\hat{\sigma}^2)$ proposed in Eq. (6).

With regard to nonnormal distributions, the results depend on its kurtosis level. The confidence intervals resulting from Eq. (5) are slightly conservative in platykurtic distributions and slightly liberal in moderately leptokurtic or skewed distributions. With highly nonnormal distributions the coverage probability of (5) can be considerably less than $1-\alpha$ unless n is large. In contrast to (5), (1) is very conservative in platykurtic distributions, very liberal in leptokurtic distributions, and its coverage probability does not converge to $1-\alpha$ as n increases. Clearly (5) is superior to (1) for all distributions considered in Tables 3 and 4. When compared to the standard error *se*(1) (Table 3), the estimated coverage probability improves when the alternative standard error for the $\ln(\hat{\sigma}^2)$: *se*(2), is used (Table 4). From the three estimators of γ_4, the one based on median, $\hat{\gamma}_4(3)$, results in confidence intervals for σ^2 with estimated coverage probability closer to the true confidence level of 95%. Thus, confidence intervals for the population variance resulting from (5) are superior to the ones resulting from (1) and the standard error of $\ln(\hat{\sigma}^2)$ improves when the median is used to estimate γ_4.

3. Confidence interval for the ratio of variances

For two independent random variables with $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$, the confidence interval for the ratio of population variances is:

Table 4. Estimated 95% probabilities of C. I. in (5) considering standard error *se*(2).

Dist.	n	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$	Dist.	n	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
N(0, 1)	10	0.950	0.956	0.969	Lap	10	0.934	0.943	0.957
	20	0.941	0.943	0.956		20	0.924	0.927	0.940
	30	0.942	0.944	0.954		30	0.927	0.930	0.938
	40	0.943	0.944	0.952		40	0.927	0.928	0.936
	50	0.942	0.943	0.950		50	0.929	0.930	0.936
	100	0.944	0.944	0.947		100	0.936	0.936	0.940
U(0, 1)	10	0.962	0.966	0.984	t(5)	10	0.918	0.930	0.946
	20	0.947	0.949	0.974		20	0.902	0.908	0.922
	30	0.947	0.949	0.970		30	0.902	0.905	0.915
	40	0.947	0.948	0.967		40	0.901	0.904	0.912
	50	0.948	0.949	0.964		50	0.905	0.907	0.914
	100	0.948	0.948	0.958		100	0.912	0.913	0.917
B(3, 3)	10	0.957	0.961	0.977	G(1, 6)	10	0.875	0.898	0.932
	20	0.947	0.949	0.965		20	0.874	0.890	0.931
	30	0.946	0.947	0.960		30	0.884	0.897	0.940
	40	0.948	0.948	0.959		40	0.892	0.902	0.946
	50	0.948	0.948	0.957		50	0.900	0.909	0.952
	100	0.948	0.948	0.953		100	0.918	0.926	0.963
Log	10	0.934	0.943	0.957	Exp	10	0.839	0.870	0.906
	20	0.924	0.927	0.940		20	0.840	0.861	0.909
	30	0.927	0.930	0.938		30	0.853	0.869	0.919
	40	0.927	0.928	0.936		40	0.865	0.876	0.926
	50	0.929	0.930	0.936		50	0.874	0.884	0.933
	100	0.936	0.936	0.940		100	0.895	0.905	0.946
B(1, 10)	10	0.876	0.898	0.931	χ^2_1	10	0.765	0.816	0.874
	20	0.874	0.891	0.932		20	0.794	0.825	0.895
	30	0.885	0.898	0.941		30	0.815	0.837	0.907
	40	0.893	0.904	0.947		40	0.832	0.848	0.917
	50	0.902	0.910	0.953		50	0.844	0.857	0.923
	100	0.920	0.928	0.965		100	0.876	0.888	0.941

The coverage probability is computed based on 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 from various theoretical distributions: standard normal: N(0, 1), uniform: U(0, 1), beta: B(3, 3) and B(1, 10), logistic: Log, Laplace: Lap, Student's *t* with 5 degrees of freedom: t(5), gamma: G(1, 6), exponential: Exp, and chi-squared with 1 degree of freedom: χ^2_1. $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ are the estimators for the standardized fourth central moment; see Eqs. (3) and (4). The confidence intervals are computed based on *se*(2), the standard error of $\ln(\hat{\sigma}^2)$ proposed in Eq. (6).

$$\left[IC_{\frac{\sigma_2^2}{\sigma_1^2}}\right]_\lambda = \left]\frac{S_2^2}{S_1^2}\cdot f_1; \frac{S_2^2}{S_1^2}\cdot f_2\right[, \tag{7}$$

where f_1 and f_2 are the quantiles of the F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.

As $\ln\left(\frac{\sigma_2^2}{\sigma_1^2}\right) = \ln(\sigma_2^2) - \ln(\sigma_1^2)$, a confidence interval for $\frac{\sigma_2^2}{\sigma_1^2}$ can also be deduced from a reverse-transformed confidence interval of $\left[\ln(\sigma_2^2) - \ln(\sigma_1^2)\right]$:

$$\exp\left\{\left[\ln\left(c_2\hat{\sigma}_2^2\right) - \ln\left(c_1\hat{\sigma}_1^2\right)\right] \pm z_{1-\alpha/2}\sqrt{se(2)_1^2 + se(2)_2^2}\right\}. \tag{8}$$

We do not consider $se(1)$ due to the best performance of $se(2)$ in case of a single variance. To compare the coverage probability of the new confidence interval for the ratio of two populations variance resulting from Eq. (8), we consider also the confidence interval resulting from Eq. (7) – identified by "Normal" in the next table – and the confidence interval resulting from Eq. (7) but with the adjustment in degrees of freedom (r_1 and r_2) suggested by Shoemaker (2003):

$$r_i = \frac{2n_i}{\frac{\mu_4}{\sigma^4} - \frac{n_i - 3}{n_i - 1}}, \quad i = 1, 2, \tag{9}$$

where μ_4 is the fourth moment about the population mean, σ is the standard deviation and n_i is the size of sample i. Thus, we assume that $\sigma_2^2 S_1^2/\sigma_1^2 S_2^2$ has an $F-$ Snedecor distribution with r_1

Table 5. Estimated 95% probabilities of C. I. for the ratio of populations variance.

Dist.	n	Normal	F1	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$	Dist.	n	Normal	F1	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
N(0, 1)	10	0.9496	0.9931	0.9931	0.9974	0.9972	Lap	10	0.8427	0.9793	0.9793	0.9917	0.9899
	20	0.9493	0.9914	0.9914	0.9948	0.9949		20	0.8175	0.9791	0.9791	0.9868	0.9857
	30	0.9503	0.9923	0.9923	0.9944	0.9947		30	0.8093	0.9820	0.9820	0.9869	0.9863
	40	0.9499	0.9920	0.9920	0.9939	0.9943		40	0.8047	0.9835	0.9835	0.9869	0.9867
	50	0.9494	0.9928	0.9928	0.9942	0.9942		50	0.8011	0.9859	0.9859	0.9884	0.9881
	100	0.9503	0.9933	0.9933	0.9941	0.9942		100	0.7915	0.9886	0.9886	0.9896	0.9895
U(0, 1)	10	0.9886	0.9972	0.9972	0.9995	0.9991	t(5)	10	0.8740	0.9444	0.9858	0.9944	0.9938
	20	0.9946	0.9976	0.9976	0.9991	0.9991		20	0.8398	0.9498	0.9820	0.9893	0.9887
	30	0.9962	0.9975	0.9975	0.9984	0.9988		30	0.8215	0.9510	0.9829	0.9879	0.9879
	40	0.9968	0.9971	0.9971	0.9979	0.9985		40	0.8111	0.9528	0.9837	0.9875	0.9873
	50	0.9970	0.9967	0.9967	0.9978	0.9983		50	0.8050	0.9546	0.9848	0.9877	0.9877
	100	0.9977	0.9960	0.9960	0.9967	0.9973		100	0.7810	0.9546	0.9877	0.9892	0.9889
B(3, 3)	10	0.9728	0.9953	0.9953	0.9982	0.9981	G(1, 6)	10	0.8019	0.9052	0.9626	0.9976	0.9908
	20	0.9791	0.9949	0.9949	0.9971	0.9972		20	0.7479	0.9224	0.9590	0.9902	0.9888
	30	0.9819	0.9948	0.9948	0.9965	0.9969		30	0.7271	0.9314	0.9639	0.9814	0.9901
	40	0.9812	0.9945	0.9945	0.9962	0.9965		40	0.7187	0.9369	0.9698	0.9793	0.9922
	50	0.9819	0.9945	0.9945	0.9958	0.9960		50	0.7097	0.9395	0.9728	0.9781	0.9935
	100	0.9822	0.9946	0.9946	0.9955	0.9956		100	0.6935	0.9474	0.9821	0.9848	0.9959
Log	10	0.9103	0.9430	0.9890	0.9955	0.9948	Exp	10	0.7684	0.8957	0.9537	0.9969	0.9880
	20	0.8982	0.9471	0.9868	0.9920	0.9916		20	0.7317	0.9211	0.9560	0.9885	0.9879
	30	0.8939	0.9491	0.9878	0.9911	0.9907		30	0.7215	0.9310	0.9633	0.9804	0.9905
	40	0.8895	0.9500	0.9880	0.9907	0.9906		40	0.7076	0.9381	0.9695	0.9784	0.9925
	50	0.8866	0.9482	0.9886	0.9909	0.9909		50	0.7042	0.9406	0.9736	0.9790	0.9936
	100	0.8847	0.9502	0.9909	0.9920	0.9919		100	0.6902	0.9459	0.9819	0.9847	0.9959
B(1, 10)	10	0.8313	0.9038	0.9667	0.9982	0.9911	χ^2_1	10	0.595	0.8003	0.8814	0.9766	0.9652
	20	0.8102	0.9217	0.9677	0.9930	0.9915		20	0.5175	0.8074	0.8784	0.8866	0.9634
	30	0.8075	0.9311	0.9731	0.9881	0.9930		30	0.4674	0.7926	0.8781	0.8744	0.9617
	40	0.8027	0.9367	0.9775	0.9858	0.9944		40	0.4284	0.7735	0.8734	0.8682	0.9592
	50	0.8035	0.9388	0.9790	0.9849	0.9954		50	0.3913	0.747	0.8657	0.8611	0.9544
	100	0.7985	0.9451	0.9862	0.9882	0.9974		100	0.2621	0.6375	0.8190	0.8272	0.9488

The coverage probability is computed based on 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 from various theoretical distributions: standard normal: N(0, 1), uniform: U(0, 1), beta: B(3, 3) and B(1, 10), logistic: Log, Laplace: Lap, Student's t with 5 degrees of freedom: t(5), gamma: G(1, 6), exponential: Exp, and chi-squared with 1 degree of freedom: χ^2_1. $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ are the estimators for the standardized fourth central moment; see Eqs. (3) and (4). The confidence intervals are computed based on Eqs. (7) (8) and (9). Column "Normal" refers to the confidence interval assuming normality; see Eq. (7). F1 refers to the confidence interval resulting from Eq. (7) but with the adjustment in degrees of freedom (r_1 and r_2) suggested by Shoemaker (2003).

and r_2 degrees of freedom. The confidence interval resulting from Eq. (9) is named by F1 in Table 5.

Estimates of coverage probabilities of confidence intervals resulting from Eqs. (7) (8), and (7) with degrees of freedom adjustment (9), were obtained using 100,000 Monte Carlo random samples of different sizes from the statistical distributions used in the section before. For the confidence intervals resulting from (8), the standard error of $\ln(\hat{\sigma}^2)$ is estimated considering the three different estimators for γ_4. The simulation results are presented in Table 5. The results show that all the confidence intervals are conservative in platykurtic distributions. For moderate leptokurtic distributions the coverage probability of confidence intervals resulting from F1, an F distribution with Shoemaker (2003) degrees of freedom adjustment, is close to $\lambda = 1 - \alpha$. In the extreme leptokurtic distribution (χ^2_1) only the coverage probability of the confidence intervals resulting from (8), with $\hat{\gamma}_4(3)$ estimator, is close to $1 - \alpha$. All the other distributions result in very liberal confidence intervals and its coverage probability does not converge to $1 - \alpha$. Thus, for moderate leptokurtic distributions, we recommend a confidence interval for the ratio of populations variance based on $F-$ Snedecor distribution with degrees of freedom adjustment proposed by Shoemaker (2003). In case of extreme leptokurtic distributions, we recommend (8) with $\hat{\gamma}_4(3)$: the estimation of the fourth central moment μ_4 is based on the median.

4. Confidence interval for the average

As we referred before, the log is the most popular transformation for right-skewed data. The purpose of this section is to provide a confidence interval for the arithmetic average of the original data X from a reverse-transformed confidence interval for the arithmetic average of the log transformation data $Y = \ln(X)$. So, we transform the original data to achieve normality (at least as approximation) but the purpose remains the statistical inference about the arithmetic average of the original data (see Feng et al. (2013) to better understand the problem). Thus, we have to reverse the confidence interval on logs to a confidence interval in the original scale.

Let X be a random variable with arithmetic mean μ_{aX} and variance σ_X^2 and let Y be the log-transformed outcome: $Y = \log(X)$. The exponentiation of the arithmetic mean of Y:

$$\mu_{aY} = n^{-1}\sum_{i=1}^{n} Y_i = n^{-1}\sum_{i=1}^{n}\log(X_i) = n^{-1}\log\left(\prod_{i=1}^{n} X_i\right) = \log\left(\prod_{i=1}^{n} X_i\right)^{n^{-1}} \tag{10}$$

is the geometric mean of X:

$$\exp(\mu_{aY}) = \exp\left[\log\left(\prod_{i=1}^{n} X_i\right)^{n^{-1}}\right] = \left(\prod_{i=1}^{n} X_i\right)^{n^{-1}} = \mu_{gX}. \tag{11}$$

Thus, the geometric mean (μ_{gX}) of the distribution of a random variable X is the exponentiation of the arithmetic mean of the natural logarithm of X.

If $X \sim$ log-normal(μ, σ^2), then the log-transformed outcome $Y = \log(X)$ has a normal distribution with mean μ and variance $\sigma^2 : Y \sim \mathcal{N}(\mu, \sigma^2)$, and the expected value of $\exp(Y)$ is:

$$E[\exp(Y)] = \underbrace{E(X)}_{\mu_{aX}} = \exp\left(\mu + \frac{\sigma^2}{2}\right) = \exp\left(\frac{\sigma^2}{2}\right)\underbrace{\exp(\mu_{aY})}_{\mu_{gX}}, \tag{12}$$

where $\exp(\mu_{aY})$ is the geometric mean of X (see Eq. (11)). This equation shows that a simple adjustment to the geometric mean is needed to obtain the arithmetic mean of X. Because $\sigma^2 > 0$, $\exp\left(\frac{\sigma^2}{2}\right) > 1$, and for large σ^2 this adjustment factor can be substantially larger than unity.

The μ_{gX} plays an important role in a log-normal distribution because the distribution of a ratio of log-normal random variables has a known log-normal distribution, and the geometric mean of a log-normal ratio is equal to the ratio of the individual geometric means (no such convenient property holds for arithmetic means with log-normal data, with either differences or ratios).

Equation (12) relies on the normality of Y, assuming that the distribution of X is log-normal. However, right-skewed data does not imply that the data generating process is log-normal and it is useful to have an adjustment factor that does not rely on normality (Wooldridge 2020):

$$E(X) = \mu_{aX} = \gamma\exp(\mu_{aY}) = \gamma\mu_{gX}. \tag{13}$$

The $(1-\alpha)\times 100\%$ confidence interval for the exponential of the arithmetic average of Y and the geometric average of X is:

$$\left]\underbrace{\exp\left(\hat{\mu}_{aY} - t_{n-1,\alpha/2}\cdot\hat{\sigma}_Y/\sqrt{n}\right)}_{LL}; \underbrace{\exp\left(\hat{\mu}_{aY} + t_{n-1,\alpha/2}\cdot\hat{\sigma}_Y/\sqrt{n}\right)}_{UL}\right[\tag{14}$$

where $t_{n-1,\alpha/2}$ is the $(1-\alpha/2)$ quantile of Student's t distribution with $n-1$ degrees of freedom. $\hat{\mu}_{aY}$ and $\hat{\sigma}_Y^2$ are the sample arithmetic mean and variance of Y. Since the geometric mean is a monotonic function of the mean of the logarithms, the upper and lower confidence limits for the geometric mean of X are the exponential of the confidence limits for μ_{aY}. Eq. (14) relies on the normality of Y, assuming that the distribution of X is log-normal. If only the approximation to normality is feasible, we replace $t_{n_1+n_2-2,\alpha/2}$ by $z_{1-\alpha/2}$, the quantile of the standardized normal distribution. As Galton (1897) suggested in one of the earliest papers on geometric average, the distribution of $\hat{\mu}_{aY}$ will approach normality as n increases, for all parent distributions to which the central limit theorem applies. Thus, the distribution of $\hat{\mu}_{gX}$ will approach the log-normal form, even though the parent distribution of X may not be log-normal (Alf and Grossberg 1979).

The confidence interval for the arithmetic mean of X results from the product of the limits in (14) by the estimate of γ (see Eq. (13)):

$$(\hat{\gamma}LL; \hat{\gamma}UL). \tag{15}$$

To estimate γ we follow the procedure used by Wooldridge (2020) through the regression of the arithmetic average of X: $E(X) = \mu_{aX}$, on the single variable (the geometric average of $X : \mu_{gX}$), without an intercept; that is, we perform a simple regression through the origin. The coefficient on μ_{gX} is the estimate of γ. If only a sample is available, we can bootstrap the sample to get different estimates for μ_{aX} and μ_{gX} and proceed with the estimation.

To assess the coverage probability of the confidence interval in (15), identified by $\hat{\gamma}\hat{\mu}_{gX}$, we compare it (see Table 6) with confidence intervals computed in three different ways: assuming the normality of X (Normal), considering the estimate of variance proposed by Bonett (2006): $\exp\left[\ln\left(c\hat{\sigma}^2\right)\right]$, and the one resulting from the corrected t variable which is derived by using a Cornish and Fisher (1938) expansion procedure. This form for t differs from the usual variable in that the numerator is adjusted by a term involving $(\hat{\mu}_a - \mu_a)^2$ and a constant. These adjustments correct bias and skewness effects due to the skewness of the nonnormal distributions (see Johnson (1978) for details):

$$t' = \left[(\hat{\mu}_a - \mu_a) + \frac{\mu_3}{6\sigma^2 n}\right]\left(\frac{\sigma^2}{n}\right)^{-\frac{1}{2}} \tag{16}$$

and the endpoints of the resulting $(1-\alpha)$ percent confidence interval for the population arithmetic average would be:

$$\left(\hat{\mu}_a + \frac{\hat{\mu}_3}{6\hat{\sigma}^2 n}\right) \pm t_{n-1,\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}. \tag{17}$$

Estimates of the coverage probability of the four confidence intervals were obtained using 100,000 Monte Carlo random samples of different sizes: 10, 20, 30, 40, 50 and 100, from different distributions for which the logarithm of the random variable X is defined ($X > 0$). The simulation results are shown in Table 6. In case of symmetric distributions, the coverage probability of the Normal, Bonett and Johnson confidence intervals is very similar and closer to the nominal confidence level, when compared to the one resulting from $\hat{\gamma}\widehat{GM}_X$. The conclusion is different when right-skewed distributions are considered. Thus, by log-transforming the original data, computing a confidence interval for the geometric average and then computing a confidence interval for the arithmetic average based on Eqs. (13) (14) and (15) results in a coverage probability that is very close to the nominal confidence level, no matter the sample size. Thus, our recommendation is to use this procedure to obtain confidence intervals for the arithmetic average in case of right-skewed distributions.

5. Confidence interval for the difference of averages

In this section we compare the coverage probability of four confidence intervals for the difference of the arithmetic averages of two populations. All the confidence intervals are derived from the results of Sec. 4 for the univariate case. The first one, that we identify by "Normal", is given by:

Table 6. Estimated 95% probabilities of C. I. for arithmetic average.

Dist.	n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$	Dist.	n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$
U(0,1)	10	0.9460	0.9626	0.9475	0.9295	W(1,1)	10	0.8998	0.9195	0.9045	0.9499
	20	0.9473	0.9578	0.9486	0.9369		20	0.9185	0.9288	0.9201	0.9486
	30	0.9509	0.9577	0.9491	0.9410		30	0.9266	0.9333	0.9293	0.9491
	40	0.9495	0.9548	0.9505	0.9420		40	0.9312	0.9366	0.9333	0.9488
	50	0.9496	0.9537	0.9499	0.9427		50	0.9354	0.9398	0.9367	0.9502
	100	0.9505	0.9527	0.9498	0.9471		100	0.9416	0.9440	0.9430	0.9498
B(3,3)	10	0.9482	0.9649	0.9487	0.9379	G(1,6)	10	0.8379	0.8608	0.9120	0.9481
	20	0.9489	0.9593	0.9496	0.9423		20	0.8680	0.8794	0.9237	0.9479
	30	0.9484	0.9558	0.9505	0.9433		30	0.8813	0.8882	0.9290	0.9491
	40	0.9501	0.9553	0.9511	0.9457		40	0.8914	0.8969	0.9344	0.9494
	50	0.9495	0.9542	0.9496	0.9463		50	0.8983	0.9027	0.9362	0.9491
	100	0.9504	0.9528	0.9515	0.9490		100	0.9169	0.9192	0.9441	0.9492
LN(0,1)	10	0.8379	0.8608	0.8431	0.9481	Exp	10	0.8993	0.9183	0.9036	0.9501
	20	0.8680	0.8794	0.8718	0.9479		20	0.9191	0.9291	0.9217	0.9495
	30	0.8813	0.8882	0.8861	0.9491		30	0.9277	0.9345	0.9304	0.9483
	40	0.8914	0.8969	0.8971	0.9494		40	0.9323	0.9378	0.9335	0.9501
	50	0.8983	0.9027	0.9029	0.9491		50	0.9343	0.9384	0.9361	0.9490
	100	0.9169	0.9192	0.9194	0.9492		100	0.9424	0.9444	0.9421	0.9486
B(1,10)	10	0.9134	0.9317	0.9155	0.9489	χ_1^2	10	0.8612	0.8807	0.8631	0.9487
	20	0.9277	0.9374	0.9311	0.9488		20	0.8908	0.9006	0.8983	0.9486
	30	0.9342	0.9415	0.9361	0.9486		30	0.9078	0.9150	0.9104	0.9475
	40	0.9382	0.9436	0.9393	0.9484		40	0.9176	0.9230	0.9198	0.9488
	50	0.9411	0.9453	0.9431	0.9482		50	0.9231	0.9273	0.9250	0.9493
	100	0.9437	0.9457	0.9459	0.9480		100	0.9340	0.9363	0.9366	0.9489

We simulate 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 from distributions: uniform: U(0,1), beta: B(3,3) and B(1,10), log-normal: LN(0,1), Weibull: W(1,1), Gamma: G(1,6), exponential: Exp, and chi-squared with 1 degree of freedom: χ_1^2. Normal: $\hat{\mu}\pm t_{n-1,1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$. Bonett: $\hat{\mu}\pm z_{1-\alpha/2}\frac{\hat{\sigma}_B}{\sqrt{n}}$, where $\hat{\sigma}_B^2=\exp\left[\ln\left(c\hat{\sigma}^2\right)\right]$, see Eq. (5). Johnson: $\left(\hat{\mu}+\frac{\hat{\mu}_3}{6\hat{\sigma}^2 n}\right)\pm t_{n-1,1-\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$, see Eqs. (16) and (17). $\hat{\gamma}\widehat{GM}_X$: $(\hat{\gamma}LL;\hat{\gamma}UL)$, see Eqs. (14) and (15).

$$(\hat{\mu}_1-\hat{\mu}_2)\pm z_{1-\alpha/2}\sqrt{\frac{\hat{\sigma}_1^2}{n_1}+\frac{\hat{\sigma}_2^2}{n_2}}, \tag{18}$$

where $z_{1-\alpha/2}$ is the quantile of the standardized normal distribution, n_i, $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ are the size, the average and the variance of sample i $(i=1,2)$, respectively.

The second one is based on the variance adjustment proposed by Bonett (2006):

$$(\hat{\mu}_1-\hat{\mu}_2)\pm z_{1-\alpha/2}\sqrt{\frac{\hat{\sigma}_{B1}^2}{n_1}+\frac{\hat{\sigma}_{B2}^2}{n_2}}, \tag{19}$$

where $\hat{\sigma}_{Bi}^2=\exp\left[\ln\left(c_i\hat{\sigma}_i^2\right)\right]$, see Eq. (5), and $c_i=n_i/(n_i-z_{1-\alpha/2})$.

The third confidence interval is based on the result of Johnson (1978):

$$\left(\bar{\mu}_1+\frac{\hat{\mu}_{1,3}}{6\hat{\sigma}_1^2 n_1}\right)-\left(\bar{X}_2+\frac{\hat{\mu}_{2,3}}{6\hat{\sigma}_1^2 n_2}\right)\pm t_{n_1+n_2-2,\alpha/2}\sqrt{\frac{\hat{\sigma}_1^2}{n_1}+\frac{\hat{\sigma}_2^2}{n_2}}. \tag{20}$$

As the geometric averages of the original data X_i are given by: $\exp\left(\mu_{Y_1}\right)$ and $\exp\left(\mu_{Y_2}\right)$, and

$$\frac{\mu_{X_1}}{\mu_{X_2}}=\frac{\gamma_1\exp\left(\mu_{Y_1}\right)}{\gamma_2\exp\left(\mu_{Y_2}\right)}=\frac{\gamma_1}{\gamma_2}\exp\left(\mu_{Y_1}-\mu_{Y_2}\right),$$

the fourth confidence interval for the ratio of two arithmetic averages is derived from the confidence interval for the ratio of two geometric averages:

Table 7. Estimated 95% probabilities of C. I. for the difference of arithmetic averages.

Dist.	n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$	Dist.	n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$
U(0,1)	10	0.9477	0.9626	0.9395	0.9360	W(1,1)	10	0.9476	0.9647	0.9496	0.9372
	20	0.9488	0.9593	0.9459	0.9455		20	0.9498	0.9595	0.9512	0.9441
	30	0.9495	0.9557	0.9455	0.9452		30	0.9496	0.9561	0.9503	0.9461
	40	0.9507	0.9546	0.9473	0.9471		40	0.9491	0.9547	0.9498	0.9460
	50	0.9505	0.9523	0.9482	0.9462		50	0.9491	0.9533	0.9495	0.9470
	100	0.9506	0.9521	0.9494	0.9478		100	0.9499	0.9518	0.9500	0.9474
B(3,3)	10	0.9238	0.9459	0.9257	0.9089	G(1,6)	10	0.9488	0.9655	0.9487	0.9378
	20	0.9368	0.9481	0.9378	0.9303		20	0.9495	0.9587	0.9490	0.9437
	30	0.9407	0.9476	0.9413	0.9353		30	0.9503	0.9569	0.9503	0.9459
	40	0.9443	0.9484	0.9441	0.9391		40	0.9502	0.9551	0.9499	0.9473
	50	0.9440	0.9480	0.9443	0.9411		50	0.9500	0.9543	0.9505	0.9474
	100	0.9443	0.9472	0.9452	0.9421		100	0.9511	0.9532	0.9504	0.9483
LN(0,1)	10	0.9604	0.9746	0.9479	0.9518	Exp	10	0.9517	0.9681	0.9506	0.9436
	20	0.9601	0.9676	0.9489	0.9513		20	0.9518	0.9616	0.9509	0.9465
	30	0.9573	0.9649	0.9496	0.9529		30	0.9514	0.9584	0.9511	0.9477
	40	0.9571	0.9629	0.9503	0.9522		40	0.9509	0.9560	0.9508	0.9475
	50	0.9565	0.9606	0.9501	0.9514		50	0.9515	0.9556	0.9513	0.9487
	100	0.9538	0.9553	0.9500	0.9480		100	0.9507	0.9529	0.9502	0.9481
B(1,10)	10	0.9258	0.9488	0.9273	0.9160	χ_1^2	10	0.9546	0.9711	0.9511	0.9479
	20	0.9386	0.9499	0.9396	0.9325		20	0.9538	0.9637	0.9514	0.9481
	30	0.9412	0.9487	0.9425	0.9370		30	0.9534	0.9601	0.9505	0.9488
	40	0.9424	0.9481	0.9424	0.9393		40	0.9529	0.9581	0.9507	0.9492
	50	0.9450	0.9497	0.9453	0.9422		50	0.9533	0.9576	0.9506	0.9500
	100	0.9445	0.9468	0.9457	0.9417		100	0.9496	0.9520	0.9494	0.9465

We simulate 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 (population 1) and 15, 25, 35, 45, 55, 110 (population 2) from distributions: uniform: U(0,1), beta: B(3,3) and B(1,10), log-normal: LN(0,1), Weibull: W(1,1), Gamma: G(1,6), exponential: Exp, and chi-squared with 1 degree of freedom: χ_1^2. Normal: see Eq. (18), Bonett: see Eq. (19), Johnson: see Eq. (20) and $\hat{\gamma}\widehat{GM}_X$: see Eqs. (21) and (22).

$$\left]\underbrace{\exp\left(\hat{\mu}_{Y_1}-\hat{\mu}_{Y_2}-t_{n_1+n_2-2,\alpha/2}\sqrt{\frac{\hat{\sigma}^2_{Y_1}}{n_1}+\frac{\hat{\sigma}^2_{Y_2}}{n_2}}\right)}_{LL};\underbrace{\exp\left(\hat{\mu}_{Y_1}-\hat{\mu}_{Y_2}+t_{n_1+n_2-2,\alpha/2}\sqrt{\frac{\hat{\sigma}^2_{Y_1}}{n_1}+\frac{\hat{\sigma}^2_{Y_2}}{n_2}}\right)}_{UL}\right[\tag{21}$$

where $n_1 + n_2 - 2$ is the $(1-\alpha/2)$ quantile of Student's t distribution with $n_1 + n_2 - 2$ degrees of freedom. $\hat{\mu}_{Y_i}$ and $\hat{\sigma}^2_{Y_i}$ are the sample mean and variance of Y_i. Eq. (21) relies on the normality of Y_i, assuming that the distribution of X_i is log-normal. If only the approximation to normality is feasible, we replace $t_{n_1+n_2-2,\alpha/2}$ by $z_{1-\alpha/2}$, the quantile of the standardized normal distribution.

After computing the confidence limits in (21) it is possible to obtain a confidence interval for the ratio of two arithmetic averages, which results from the product of the limits by the ratio of the estimates for γ (see Eq. (13)):

$$\left(\frac{\hat{\gamma}_1}{\hat{\gamma}_2}\cdot LL;\frac{\hat{\gamma}_1}{\hat{\gamma}_2}\cdot UL\right). \tag{22}$$

Estimates of the coverage probability of the four confidence intervals were obtained using 100,000 Monte Carlo random samples of different sizes: 10, 20, 30, 40, 50 and 100 (population 1) and 15, 25, 35, 45, 55, 110 (population 2), from different distributions (see Table 7). As one can see, the confidence intervals give approximately the same coverage and they are very close to the nominal confidence level of 95%. However, $\hat{\gamma}\widehat{GM}_X$, the confidence interval resulting from Eqs. (21) and (22) is the one with the smallest length (see Table 8), providing more efficient estimates.

Table 8. Length of the C. I. for the difference of arithmetic averages.

Dist.	n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$	Dist.	n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$
U(0,1)	10	0.4848	0.5324	1.7984	0.4594	W(1,1)	10	0.7710	0.8465	1.1309	0.7308
	20	0.3484	0.3651	1.2609	0.3385		20	0.5565	0.5832	0.7960	0.5407
	30	0.2867	0.2958	1.0256	0.2810		30	0.4586	0.4733	0.6501	0.4497
	40	0.2492	0.2552	0.8863	0.2455		40	0.3991	0.4087	0.5621	0.3932
	50	0.2235	0.2279	0.7908	0.2208		50	0.3581	0.3650	0.5033	0.3538
	100	0.1572	0.1587	0.5503	0.1553		100	0.2520	0.2544	0.3513	0.2490
B(3,3)	10	0.3158	0.3468	0.7915	0.2992	G(1,6)	10	0.6773	0.7436	0.7318	0.6421
	20	0.2275	0.2384	0.5639	0.2210		20	0.4892	0.5126	0.5206	0.4753
	30	0.1873	0.1933	0.4623	0.1836		30	0.4036	0.4165	0.4265	0.3957
	40	0.1630	0.1669	0.4016	0.1606		40	0.3516	0.3600	0.3705	0.3463
	50	0.1462	0.1490	0.3596	0.1444		50	0.3152	0.3213	0.3316	0.3114
	100	0.1029	0.1038	0.2524	0.1016		100	0.2220	0.2241	0.2326	0.2193
LN(0,1)	10	3.1637	3.4695	2.0164	3.0051	Exp	10	1.6231	1.7814	2.8072	1.5396
	20	2.3708	2.4840	1.3328	2.3046		20	1.1851	1.2419	1.7959	1.1516
	30	1.9919	2.0555	1.0612	1.9535		30	0.9799	1.0112	1.4111	0.9608
	40	1.7607	1.8030	0.9102	1.7346		40	0.8547	0.8753	1.2033	0.8420
	50	1.5882	1.6188	0.8076	1.5692		50	0.7678	0.7826	1.0587	0.7586
	100	1.1402	1.1511	0.5559	1.1265		100	0.5423	0.5474	0.7209	0.5358
B(1,10)	10	0.1361	0.1494	2.8000	0.1290	χ^2_1	10	2.2419	2.4598	9.2753	2.1277
	20	0.0988	0.1035	1.7477	0.0960		20	1.6529	1.7319	4.3789	1.6063
	30	0.0817	0.0843	1.3699	0.0801		30	1.3730	1.4169	3.0963	1.3463
	40	0.0712	0.0729	1.1594	0.0701		40	1.1994	1.2283	2.4761	1.1816
	50	0.0639	0.0652	1.0267	0.0632		50	1.0781	1.0989	2.1196	1.0652
	100	0.0451	0.0455	0.6995	0.0446		100	0.7641	0.7714	1.3402	0.7549

We simulate 100,000 Monte Carlo samples of different sizes: 10, 20, 30, 40, 50 and 100 (population 1) and 15, 25, 35, 45, 55, 110 (population 2) from distributions: uniform: U(0,1), beta: B(3,3) and B(1,10), log-normal: LN(0,1), Weibull: W(1,1), Gamma: G(1,6), exponential: Exp, and chi-squared with 1 degree of freedom: χ^2_1. Normal: see Eq. (18), Bonett: see Eq. (19), Johnson: see Eq. (20) and $\hat{\gamma}\widehat{GM}_X$: see Eqs. (21) and (22).

Table 9. Descriptive statistics.

Statistics	Payables	Receivables
Arithmetic mean	11317.68	11472.53
Geometric mean	2472.40	3337.83
Harmonic mean	1075.43	1783.39
Standard Error	320.10	339.51
Median	1650.00	2040.00
Mode	516.60	1600.00
Standard Deviation	25344.79	25333.77
Sample Variance	642358281.99	641799744.53
Kurtosis	17.86	19.23
Skewness	3.63	3.79
Range	196919.11	195196.49
Minimum	400.53	600.22
Maximum	197319.64	195796.71
Sum	70950545.59	63879033.09
Count	6269	5568
Confidence Level(95.0%)	627.51	665.57

6. Application to accounts payable and receivable data sets

Populations in auditing and accounting are almost always skewed to the right: values are often very low, but are occasionally high or very high. In this empirical application (with real data) we use two populations composed by the accounts payable and receivable of a Portuguese company during the year 2019. The empirical study is conducted in this way. First, we extract random samples with different sizes: 10, 20, 30, 40, 50, 100, from the populations of accounts payable and receivable. Then we compare the coverage probability of the confidence intervals derived in the sections before. The empirical results confirm the superiority of these intervals, already noticed in the Monte Carlo simulation studies.

Table 10. Bias of $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$.

	Payables			Receivables		
n	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
10	−15.866	−15.325	−10.425	−15.849	−14.571	−11.508
20	−8.419	−7.399	−3.483	−11.717	−10.711	−6.839
30	−7.628	−6.769	−2.750	−11.180	−10.315	−5.616
40	−6.532	−5.529	−1.793	−10.086	−8.021	−4.656
50	−6.235	−4.550	−1.246	−6.755	−5.702	−1.922
100	−5.550	−3.450	−0.515	−3.920	−2.549	−0.679

$\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ are the estimators for the standardized fourth central moment; see Eqs. (3) and (4). The bias is computed as $\hat{\gamma}_4(j) - \gamma_4(j)$, where $\gamma_4(j)$ is the true value of kurtosis.

Table 11. Estimated 95% probabilities of C. I. for the population variance.

	Payables				Receivables			
n	Normal	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$	Normal	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
10	0.434	0.619	0.687	0.762	0.442	0.612	0.672	0.735
20	0.370	0.737	0.770	0.833	0.390	0.734	0.771	0.831
30	0.333	0.799	0.820	0.875	0.352	0.782	0.806	0.866
40	0.318	0.834	0.852	0.903	0.337	0.814	0.833	0.886
50	0.310	0.854	0.868	0.919	0.326	0.840	0.853	0.902
100	0.286	0.906	0.918	0.952	0.294	0.900	0.913	0.946

The column "Normal" refers to the coverage probability of the confidence interval assuming the normality of payables and receivables; see Eq. (1). $\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ represent the coverage probabilities of the confidence intervals resulting from Eq. (5) considering the three estimators for the fourth central moment μ_4 in *se*(2), the standard error of the $\ln(\hat{\sigma}^2)$.

Table 12. Estimated 95% probabilities of C. I. for the ratio of populations variance.

n	Normal	F1	$\hat{\gamma}_4(1)$	$\hat{\gamma}_4(2)$	$\hat{\gamma}_4(3)$
10	0.398	0.762	0.756	0.913	0.89335
20	0.419	0.855	0.847	0.886	0.93107
30	0.439	0.888	0.898	0.906	0.95407
40	0.451	0.903	0.924	0.926	0.96598
50	0.459	0.913	0.941	0.942	0.97507
100	0.481	0.934	0.974	0.976	0.99048

$\hat{\gamma}_4(1), \hat{\gamma}_4(2)$ and $\hat{\gamma}_4(3)$ are the estimators for the standardized fourth central moment; see Eqs. (3) and (4). The confidence intervals are computed based on Eqs. (7) (8) and (9). Column "Normal" refers to the confidence interval assuming normality; see Eq. (7). F1 refers to the confidence interval resulting from Eq. (7) but with the adjustment in degrees of freedom suggested by Shoemaker (2003).

Descriptive statistics are shown in Table 9. The dimension of the accounts payable and receivable populations is 6269 and 5568, respectively. The distributions are right-skewed and highly leptokurtic. Thus, we expect that the coverage probability of the confidence intervals proposed in this paper should be superior when compared to the existing ones.

First we analyze the bias of the three estimators of the kurtosis by comparing the Pearson estimator with the alternative estimators in Eq. (4). 100,000 random samples are selected with different sizes for accounts payable and receivable and the average of bias per estimator is shown in Table 10. The empirical results confirm the Monte Carlo simulation study based on several distributions (see Table 1). The bias of the estimators is almost always negative and the estimator $\hat{\gamma}_4(3)$, where the central tendency is represented by the median, is the one with the smallest bias confirming its appropriateness to estimate the kurtosis of leptokurtic and/or right-skewed empirical distributions.

Next we compute confidence intervals for the variance of accounts payable and receivable based on Eqs. (1) and (5), and in the last case the standard error of $\ln(\hat{\sigma}^2)$ is computed based on ($se(2)$ of Eq. (6) and considering the three different estimators for kurtosis. The results are presented in Table 11. As one can see, the confidence interval assuming the normality of the

Table 13. Estimated 95% probabilities of C. I. for the arithmetic average.

	Payables				Receivables			
n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$
10	0.725	0.743	0.731	0.906	0.707	0.725	0.712	0.906
20	0.816	0.825	0.822	0.931	0.807	0.815	0.813	0.930
30	0.857	0.863	0.863	0.937	0.849	0.855	0.855	0.938
40	0.881	0.886	0.887	0.941	0.872	0.877	0.878	0.942
50	0.893	0.897	0.898	0.942	0.889	0.893	0.894	0.944
100	0.923	0.925	0.927	0.948	0.919	0.922	0.924	0.949

Normal: $\bar{X} \pm t_{n-1,\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$. Bonett: $\bar{X} \pm t_{n-1,\alpha/2}\frac{\hat{\sigma}_B}{\sqrt{n}}$, where $\hat{\sigma}_B^2 = \exp\left[\ln\left(c\hat{\sigma}^2\right)\right]$, see Eq. (5). Johnson: $\left(\bar{X} + \frac{\hat{\mu}_3}{6\hat{\sigma}^2 n}\right) \pm t_{n-1,\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$, see Eqs. (16) and (17). $\hat{\gamma}\widehat{GM}_X$: $(\hat{\gamma}LL; \hat{\gamma}UL)$, see Eq. (15).

Table 14. Estimated 95% probabilities of C. I. for the ratio of arithmetic averages.

n	Normal	Bonett	Johnson	$\hat{\gamma}\widehat{GM}_X$
10	0.982	0.993	0.964	0.947
20	0.968	0.978	0.953	0.948
30	0.960	0.968	0.949	0.949
40	0.957	0.963	0.948	0.950
50	0.955	0.959	0.948	0.950
100	0.953	0.955	0.950	0.951

Normal: see Eq. (18), Bonett: see Eq. (19), Johnson: see Eq. (20) and $\hat{\gamma}\widehat{GM}_X$: see Eqs. (21) and (22).

populations ("Normal") is very liberal and its coverage probability does not converge to the 95% confidence level when the sample size increases. The confidence intervals resulting from Eq. (5) seem slightly liberal for small to moderate samples. However, when the sample size increases to 100, the coverage probability of the confidence interval resulting from $se(2)$, with $\hat{\gamma}_4(3)$ as the estimator of kurtosis, is very close to the nominal confidence level. Thus, the empirical results support the Monte Carlo simulation study (see Table 3) and highlights the superiority of (5) when compared to (1) in case of leptokurtic distributions.

The next step is to compute the confidence interval for the ratio of two populations variance. The results are shown in Table 12. Due to the asymmetry and excess of kurtosis of the accounts payable and receivable distributions, the coverage probability of the confidence interval assuming the normality (column "Normal") is lower and very distant from the nominal confidence level. The coverage probability of confidence intervals resulting from F1, an F distribution with Shoemaker (2003) degrees of freedom adjustment, is close to $\lambda = 1 - \alpha$ for big sample sizes. The confidence intervals resulting from (8) are preferable for small to moderate sample sizes. The coverage probability of the confidence interval resulting from (8) with $\hat{\gamma}_4(3)$, where the estimation of the fourth central moment μ_4 is based on the median, is very close to the nominal confidence level. Thus, it seems more appropriate in case of moderate sample sizes. The confidence interval resulting from (8) and based on $\hat{\gamma}_4(2)$ seems more appropriate for small sample sizes.

Next we compute and compare the coverage probability of confidence intervals for the population arithmetic average of accounts payable and receivable (see Table 13). The results confirm the conclusions of the Monte Carlo simulations studies of Sec. 4. By log-transforming the original data, computing a confidence interval for the geometric average and then computing a confidence interval for the arithmetic average of original data based on Eqs. (13) (14) and (15) results in a coverage probability that is very close to the nominal confidence level, no matter the sample size (except when $n = 10$). Thus, the empirical result, based on real data, reinforces the recommendation to use this procedure to compute confidence intervals for the arithmetic average in case of right-skewed and/or leptokurtic distributions.

Finally we analyze the coverage probability of the confidence intervals for the ratio of two arithmetic averages of accounts payable and receivable (see Table 14). The confidence intervals, assuming the normality of the original values of accounts payable and receivable and based on the corrections proposed by Bonett (2006) and Johnson (1978), are very conservative for small samples size. For moderate to large samples size the coverage probability is approximately the same no matter is the confidence interval. After all, despite the difference being small, the confidence interval that we propose based on log-transformation (see Eqs. (21) and (22)) produces the closer coverage probability to the nominal confidence level. Thus, the empirical results confirm the Monte carlo study pointing for the usefulness of confidence interval resulting from Eqs. (21) and (22) in case of right-skewed and/or leptokurtic distributions.

7. Conclusions

In this paper we propose new confidence intervals for the population mean and variance, the ratio of two populations variance and the difference and ratio of the arithmetic averages of two populations with nonnormal distribution.

To compare the coverage probability of different confidence intervals, several Monte Carlo simulation studies have been conducted. We simulate 100,000 samples of different sizes: 10, 20, 30, 40, 50 and 100 from various theoretical distributions: standard normal, uniform, beta, logistic, Laplace, Student's t, log-normal, gamma, Weibull, exponential and chi-squared. The simulation routines have been programmed in R.

A new estimator of kurtosis based on median has been proposed, and we compare it with Pearson and another estimator based on trimmed mean suggested by Bonett (2006). We conclude that the bias of the three estimators is negative in leptokurtic distributions, understating the true value of kurtosis. The Pearson estimator, the most popular, has the largest negative bias. From the two estimators based on trimmed mean and median, the one based on median has, on average, the smallest negative bias. Thus, it seems the most appropriate to estimate the kurtosis of leptokurtic distributions.

We derive a new confidence interval for the variance of population, based on the method proposed by Bonett (2006) and considering the new estimator for kurtosis. The estimated coverage probabilities are very close to the nominal confidence level pointing for its superiority when compared to the existing methods.

We generalize this method for one single population variance to the ratio of two populations variance and we compare the resulting confidence intervals with the ones based on $F-$ Snedecor distribution with the adjustment in degrees of freedom suggested by Shoemaker (2003). The results show that all the confidence intervals are conservative in platykurtic distributions. For moderate leptokurtic distributions the coverage probability of confidence intervals resulting from the adjustment is closer to the confidence level. In the extreme leptokurtic distribution (χ_1^2) only the coverage probability of the confidence intervals resulting from the new method (based on the median to estimate the fourth central moment) is close to the confidence level.

A new confidence interval is also proposed for the arithmetic average of the original data from a reverse-transformed confidence interval for the arithmetic average of the log transformation data. Thus, we have to reverse the confidence interval on logs to a confidence interval in the original scale. When right-skewed distributions are considered, computing a confidence interval for the geometric average and then computing a confidence interval for the arithmetic average results in a coverage probability that is very close to the nominal confidence level, no matter the sample size.

Confidence intervals for the difference and the ratio of two arithmetic averages are also derived. The simulation results favor the new method only in terms of length. The confidence intervals give approximately the same coverage and they all are very close to the nominal confidence level.

Finally, we analyze two populations of accounts payable and receivable of a Portuguese company for the year 2019. The empirical results confirm the Monte Carlo simulation studies, highlighting the superiority of the new proposed methods.

Acknowledgements

I would like to thank the Editor-in-Chief Narayanaswamy Balakrishnan and the anonymous referee for their suggestions and comments.

Funding

This work was supported by Fundação para a Ciência e a Tecnologia, Grant UIDB/00315/2020.

ORCID

José Dias Curto http://orcid.org/0000-0003-2012-9015

References

Alf, E. F., and J. M. Grossberg. 1979. The geometric mean: Confidence limits and significance tests. *Perception & Psychophysics* 26 (5):419–21. doi:10.3758/BF03204171.

Atkinson, A. C. 1986. *Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis, Oxford Statistical Science Series*, 1st ed. Oxford: Oxford University Press.

Bonett, D. G. 2006. Approximate confidence interval for standard deviation of nonnormal distributions. *Computational Statistics & Data Analysis* 50 (3):775–82. doi:10.1016/j.csda.2004.10.003.

Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)* 26 (2):211–43. doi:10.1111/j.2517-6161.1964.tb00553.x.

Cahoy, D. O. 2010. A bootstrap test for equality of variances. *Computational Statistics & Data Analysis* 54 (10): 2306–16. doi:10.1016/j.csda.2010.04.012.

Chen, L. 1995. Testing the mean of skewed distributions. *Journal of the American Statistical Association* 90 (430): 767–72. doi:10.1080/01621459.1995.10476571.

Chen, H. J., and S.-Y. Chen. 1999. A nearly optimal confidence interval for the largest normal mean. *Communications in Statistics - Simulation and Computation* 28 (1):131–46. doi:10.1080/03610919908813539.

Cojbasic, V., and A. Tomovic. 2007. Nonparametric confidence intervals for population variance of one sample and the difference of variances of two samples. *Computational Statistics & Data Analysis* 51 (12):5562–78. doi: 10.1016/j.csda.2007.03.023.

Cornish, E. A., and R. A. Fisher. 1938. Moments and cumulants in the specifications of distributions. *Revue de L'institut International de Statistique / Review of the International Statistical Institute* 5 (4):307–27. doi:10.2307/1400905.

Feng, C., H. Wang, N. Lu, and X. M. Tu. 2013. Log transformation: Application and interpretation in biomedical research. *Statistics in Medicine* 32 (2):230–9. doi:10.1002/sim.5486.

Galton, F. 1897. The geometric mean in vital and social statistics. *Proceedings of the Royal Society of London* 29: 365–367.

González-Manteiga, W., R. Cao, and J. S. Marron. 1996. Bootstrap selection of the smoothing parameter in nonparametric hazard rate estimation. *Journal of the American Statistical Association* 91 (435):1130–40.

Johnson, N. J. 1978. Modified t tests and confidence intervals for asymmetrical populations. *Journal of the American Statistical Association* 73 (363):536–44. doi:10.1080/01621459.1978.10480051.

Kim, T.-H., and H. White. 2004. On robust estimation of skewness and kurtosis. *Finance Research Letters* 1 (1): 56–73. doi:10.1016/S1544-6123(03)00003-5.

Mood, A. M., F. A. Graybill, and D. C. Boes. 2007. *Introduction to the theory of statistics*. 3rd ed. McGraw-Hill, New York.

Shoemaker, L. H. 2003. Fixing the F Test for Equal Variances. *The American Statistician* 57 (2):105–14. doi:10.1198/0003130031441.

Tan, M., and L. J. Gleser. 1993. Improved point and confidence interval estimators of mean response in simulation when control variates are used. *Communications in Statistics - Simulation and Computation* 22 (4):1211–20. doi: 10.1080/03610919308813150.

Wooldridge, J. 2020. *Introductory econometrics: A modern approach*, 7th ed. South-Western, Mason.