ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

DOI: 10.5336/biostatic.2017-57348

Interval Estimation for Nonnormal Population Variance with Kurtosis Coefficient Based on Trimmed Mean

Budanmış Ortalamaya Dayalı Basıklık Katsayısı ile Normal Dağılımlı Olmayan Yığın Varyansı İçin Aralık Tahmini

Hayriye Esra AKYÜZ,[a]
Hamza GAMGAM[b]

[a]Department of Statistics,
Bitlis Eren University
Faculty of Science and Arts,
Bitlis
[b]Department of Statistics,
Gazi University Faculty of Science,
Ankara

Geliş Tarihi/*Received:* 21.07.2017
Kabul Tarihi/*Accepted:* 22.09.2017

Yazışma Adresi/*Correspondence:*
Hayriye Esra AKYÜZ
Bitlis Eren University
Faculty of Sciece and Arts,
Department of Statistics, Bitlis,
TURKEY/TÜRKİYE
heakyuz@beu.edu.tr

This paper was presented as oral presentation in the International Conference on Mathematics and Mathematics Education (ICMME-11-13 May 2017).

ABSTRACT Objective: In random experiments, most analyses are based on interpretation of the difference between the means of experiment and control groups. Therefore, studying the variance of the experiment and control groups may also be useful in interpreting the analysis results. This study focuses on interval estimation based on minimum mean-squared error biased estimators for nonnormal population variance. It is aimed that kurtosis coefficient based on trimmed mean is used instead of sample kurtosis coefficient in this study. **Material and Methods:** With the simulation study in MATLAB R2009a, confidence intervals are obtained for the estimator which is obtained using both kurtosis coefficients. Coverage probabilities and average length widths of these confidence intervals were compared. **Results:** It was determined that the coverage probabilities of the estimator which was obtained by using the kurtosis coefficient based on trimmed mean is very close to the nominal confidence level in all sample sizes. It was also observed that the average length widths which were obtained by using the kurtosis coefficient based on trimmed mean are narrower than the average widths obtained with biased sample kurtosis coefficient. It was determined that the results are the same when Type I error is different. Using sample kurtosis coefficient for nonnormal populations resulted in obtaining low coverage probabilities. **Conclusion:** It will be appropriate to prefer interval estimations based on estimator obtained with kurtosis coefficient based on Trimmed Mean since it both provides high coverage probability and narrower confidence interval for nonnormal population variance.

Keywords: Coverage probability; interval estimation; kurtosis coefficient; minimum mean-squared error; trimmed mean

ÖZET Amaç: Rasgele deneylerde bir çok analiz deney ve kontrol gruplarının ortalamaları arasındaki farkın yorumlanmasına dayalıdır. Varyansların incelenmesi de analiz sonuçlarını yorumlamada faydalı olabilir. Bu çalışma normal dağılımlı olmayan yığın varyansı için minumum hata kareler ortalamalı yanlı tahmin ediciye dayalı aralık tahmini üzerinedir. Bu çalışmada örnek basıklık katsayısı yerine budanmış ortalamaya dayalı olarak elde edilen basıklık katsayısının kullanılması amaçlanmıştır. **Gereç ve Yöntemler:** MATLAB R2009a programında yapılan simülasyon çalışmasında her iki basıklık katsayısının kullanılması ile elde edilen tahmin ediciye dayalı güven aralıkları elde edilmiştir. Bu güven aralıklarına ait kapsama olasılıkları ve ortalama aralık genişlikleri karşılaştırılmıştır. **Bulgular:** Budanmış ortalamaya dayalı basıklık katsayısının kullanılması ile elde edilen tahmin edicinin güven aralığı kapsama olasılıklarının örnek hacimlerinin tamamında nominal güven düzeyine çok yakın olduğu belirlenmiştir. Aynı zamanda budanmış ortalamaya dayalı basıklık katsayısının kullanılması sonucunda elde edilen güven aralıklarının ortalama genişliklerinin, yanlı örnek basıklık katsayısı ile elde edilen güven aralıklarının ortalama genişliklerine göre daha dar olduğu görülmüştür. Ayrıca I. Tip hata farklı iken de bu sonuçların aynı olduğu saptanmıştır. Normal dağılımlı olmayan yığınlarda kullanılan örnek basıklık katsayısının kullanılması, düşük kapsama olasılıklarının elde edilmesine sebep olmuştur. **Sonuç:** Normal dağılımlı olmayan yığın varyansı için hem yüksek kapsama olasılığı vermesi hem de daha dar bir güven aralığı vermesi nedeniyle Budanmış Ortalamaya dayalı basıklık katsayısı ile elde edilen tahmin ediciye dayalı aralık tahminlerinin tercih edilmesi uygun olur.

Anahtar Kelimeler: Kapsama olasılığı; aralık tahmini; basıklık katsayısı; minumum hata kareler ortalaması; budanmış ortalama

In randomized trials and clinical experiments with continuous outcomes, the focus of analysis is often on the mean outcomes of experimental and control groups. However, the variances of outcomes of experimental and control groups may also have a useful interpretation. Let $X_1, X_2,..., X_n$ be a random sample of size n from normal distribution. Sample variance S^2 is defined as $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, where $\bar{X}$ is sample mean. It is known that the sample variance S^2 is an unbiased estimator for population variance σ^2. An improved estimator of the variance that utilizes the kurtosis was derived by Searls and Intarapanich (1990).[1] This estimator is biased and has the minimum possible mean-squared error and it is defined as the "minimum mean-squared error biased estimator" (MBBE). The estimator has the form,

$$S_w^2 = w\sum_{i=1}^{n}(X_i - \bar{X})^2,\ w \in (0,1) \qquad (1.1)$$

where the weight is given as;

$$w = [(n+1) + (\gamma_4 - 3)n^{-1}(n-1)]^{-1} \qquad (1.2)$$

It's a function of the sample size n and the kurtosis coefficient γ_4. For all choices of weights $w \neq \frac{1}{n-1}$ the improved estimator S_w^2 is a biased estimator of population variance.[2] Note that it can express the class of estimators of the variance by,

$$S_w^2 = w(n-1)S^2, w \in \left(0, \frac{1}{n-1}\right) \qquad (1.3)$$

On the other hand expected value and mean squared error of a biased estimator S_w^2 are given as;

$$E(S_w^2) = w(n-1)\sigma^2 \qquad (1.4)$$

and

$$MSE(S_w^2) = w^2(n-1)^2 Var(S^2) + [(n-1)w - 1]^2\sigma^4 \qquad (1.5)$$

For large n_i, when random sampling from any distribution with a finite fourth moment, and by the central limit theorem, the MBBE of variance is approximately standard normal with $E(S_w^2)$ and $MSE(S_w^2)$. This estimator is always a more efficient estimator compared to unbiased estimator S^2 of the variance even in nonnormal distribution assumption.[2]

Zou et al. (2009) have suggested an interval estimation for a linear function of binomial ratios.[3] Wencheko and Chipoyera (2009) have compared the MBBE estimator S_w^2 of the population variance and unbiased estimator S^2 when kurtosis coefficient is known.[2] They have studied boxplots of different sample sizes and estimators for Normal and Exponential distribution. Longford (2010) discussed minumum mean squared error and Bayesian estimation of the variance and its common transformations in the normality and homoscedasticity with small samples.[4] Paksaranuwat and Niwitpong (2010) compared confidence intervals for the variance and the ratio of two variances when the population distributions are nonnormal and item nonresponse is occurring.[5] Burch (2011a) describe scenarios where the conventional method leads to confidence intervals that have extremely poor coverage probabilities.[6] The random variables in the model are not normally distributed Burch (2011b) constructed approximate confidence intervals based on the large-sample properties of restricted maximum likelihood estimators of variance components.[7] Donner and Zou (2012) have obtained confidence intervals for the mean and functions of standard deviation of normal distribution.[8] Niwitpong (2012) investigated the new confidence interval for the difference between two normal population variances based on the closed form method of variance estimation and derived analytic expressions of coverage probability and expected length of our proposed confidence interval compared to the existing confidence interval.[9] Rajic and Stanojevic (2013) proposed a confidence interval for the ratio of two variances based on the *t*-statistic by deriving its Edgeworth expansion and considering Hall's and Johnson's transformations.[10] Thompson (2013) derived a confidence interval for the ratio of the variances of two independent normal distributions.[11] Suwan and Niwitpong (2013) have studied interval estimation methods for a linear function of the variances of nonnormal distributions using kurtosis.[12] For the linear function of variance, coverage probabilities and average length widths of confidence intervals are studied with different sample sizes using Binom, Logit and Chi-squared distributions. They have made a comparison between interval methods where S_w^2 and S^2 are used. Banik et al. (2014) presented a simulation study which compared ten methods for constructing a confidence estimator for the population standard deviation.[13] Burch (2015) assessed the performance of approximate confidence intervals for the variance under nonnormality using modified estimators of the kurtosis.[14] Burch (2017) proposed an interval estimation procedure for the variance of a population that achieves a specified confidence level can be problematic. He determined that if the distribution of the population is known, then a distribution-dependent interval for the variance can be obtained by considering a power transformation of the sample variance.[15]

It is known that sample variance estimator S^2 does not display robust statistics features in the estimation of nonnormal population variance and the coverage probabilities of the confidence intervals obtained with this estimator have much lower values compared to the nominal confidence interval.[16,17] In such cases, it is necessary to use robust scale estimators for estimation of population variance. For nonnormal populations, sample kurtosis coefficient is a quite biased estimator.[18]

In this study, it is aimed that the kurtosis coefficient obtained with trimming proportion $0.5/\sqrt{n-4}$ which was suggested by Bonett (2006a,b) is used instead of sample kurtosis coefficient in obtaining MBBE estimator of variance.[19,20] With this information, confidence intervals based on this estimator were obtained for nonnormal population variance. Coverage probabilities and average length widths of these confidence intervals were compared with the confidence interval coverage probabilities and average length widths obtained when sample kurtosis coefficient is used.

A interval estimator was conducted in the following section for the variance of nonnormal population.

INTERVAL ESTIMATION METHOD FOR THE VARIANCE OF NONNORMAL POPULATION

In this section, interval estimator based on the MBBE estimator for the variance of a nonnormal population are included.

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n under normal distribution assumption, estimator S_w^2 is obtained as follows:

$$S_w^2 = \left(\frac{n-1}{n+1}\right) S^2 \tag{2.1}$$

In obtaining the weight coefficient given in Equation (1.2), the estimation value obtained from the sample is used instead of unknown parameter. Provided that fourth central moment is η, kurtosis coefficient is obtained as $\gamma = \frac{\eta}{\sigma^4}$.[21] This estimator is quite biased for nonnormal populations.[18] In that case, the alternative estimator suggested by Bonett (2006a,b) can be used instead of sample kurtosis coefficient for the unknown parameter.[19,20] Maximum value of this estimator is a function of sample size and this estimator is expressed as follows:[22]

$$G = n\frac{\sum_{i=1}^{n}(X_i-m)^4}{\left(\sum_{i=1}^{n}(X_i-\hat{\mu})^2\right)^2} \quad (2.2)$$

where m is the trimmed mean which was obtained with trimming proportion $0.5/\sqrt{n-4}$ and the $\hat{\mu}$ is the sample mean. While trimming is made only from the high end of the consecutive data in positively skewed distributions, trimming is made from both ends in symmetric distributions[23]. With this information, trimmed mean used for estimator G is obtained as follows in the positively skewed distributions:

$$m_1 = \frac{1}{n-u_n}\sum_{i=1}^{n-u_n} X_{(i)} \quad (2.3)$$

where $X_{(i)}$ is the ith order statistis and u_n is the number of terms to be removed from the high end of the consecutive data. If $u_n = [\rho n + 0.5]$, ρ trimming percentage and $[.]$ expression indicates the largest integer function. Trimmed mean is as follows in symmetric distributions:

$$m_2 = \frac{1}{n-2l_n}\sum_{i=l_n+1}^{n-u_n} X_{(i)} \quad (2.4)$$

In that case, S_w^2 estimator is obtained as follows for estimation of the variance of nonnormal populations:

$$S_w^2 = \frac{(n-1)}{(n+1)+(G-3)n^{-1}(n-1)}S^2 \quad (2.5)$$

It is known that estimator S_w^2 has approximately a normal distribution even when random samples are taken from any statistical distribution[2]. With this information, confidence interval for nonnormal population variance is obtained as follows using the normal distribution:

$$P\left[S_w^2 - z_{\alpha/2}\sqrt{Var(S_w^2)} \le \sigma^2 \le S_w^2 + z_{\alpha/2}\sqrt{Var(S_w^2)}\right] = 1-\alpha. \quad (2.6)$$

When the populations are nonnormal distributed, there is not a theoretical formula for $Var(S_w^2)$. For that reason, for $Var(S_w^2)$, the variance estimation value obtained from the distribution of estimator S_w^2 with Monte Carlo Simulation Method or the variance estimation value obtained with Bootstrap Method can be used.[24] For $Var(S_w^2)$, Monte Carlo Simulation Method can be expressed as follows:

Provided that the value of estimator obtained in the ith replication of the T repeated simulation study with sample data of size n is $S_{w_i}^2$, $i = 1,2,\dots,T$, the variance of the S_w^2 estimator is expressed as;

$$\widehat{Var}\,(S_w^2) = \frac{\sum_{i=1}^{T}\left(S_{w_i}^2-\bar{S}_w^2\right)^2}{T-1} \quad (2.7)$$

The $\bar{S}_w^2$ expressed in Equation (2.7) is the arithmetic mean of $S_{w_i}^2$'s. In addition, variance estimation may also be made with Bootstrap Method for $Var(S_w^2)$. For the variance of estimator S_w^2, bootstrap samples of size n and number B are generated by simple random sampling with replacement. For each Bootstrap sample, Bootstrap estimation is obtained for estimator S_w^2. This operation is repeated for B times. With the Bootstrap estimations obtained from B replications, the Bootstrap estimator for the variance of estimator S_w^2 is given as:[25]

$$\widehat{Var}\,(S_w^2) = \frac{\sum_{b=1}^{B}\left(S_{w_b}^2-\bar{S}_w^2\right)^2}{B-1} \quad (2.8)$$

Obtaining the $Var(S_w^2)$ value with both methods yields quite similar results. In the case where the population has t-distribution, variance estimation values based on 10000 replications which are obtained with Monte Carlo Simulation Method and Bootstrap Method are as follows.

TABLE 1: Variance estimation values for $Var(S_w^2)$.

	t(5)		t(10)		t(20)		t(30)	
n	MC*	Bootstrap	MC*	Bootstrap	MC*	Bootstrap	MC*	Bootstrap
10	0.8465	1.0033	0.3137	0.3336	0.2248	0.2096	0.1993	0.2337
20	0.4812	0.4534	0.1788	0.1693	0.1218	0.1199	0.1075	0.1287
30	0.4179	0.3754	0.1276	0.1250	0.0842	0.0866	0.0777	0.0756
50	0.2424	0.2442	0.0814	0.0960	0.0545	0.0536	0.0478	0.0480

*: Monte Carlo simulation method.

According to the results obtained in this table, variance estimation values based on Monte Carlo Simulation Method and Bootstrap variance estimation values have yielded quite close results in all of the sample sizes (Table 1).

SIMULATION STUDY

A simulation study was conducted with the purpose of comparing the confidence interval coverage probabilities and average length widths when sample kurtosis coefficient and the kurtosis coefficient suggested by Bonett (2006a,b)[19,20] are used. In this simulation study, the data produced from Chi-squared, Gamma, Lognormal and t-distributions with different distribution parameters was used with the program written in Matlab R2009a. In obtaining confidence intervals based on robust estimator S_w^2, trimming was made only on the high end of the consecutive data for the sample data produced from Chi-squared, Gamma and Lognormal distributions and on both ends of the consecutive data for the sample data produced from t-distribution. Using these distributions, simulation studies were conducted based on 10000 replications for $\alpha = 0.05$ and $\alpha = 0.10$ with trimming proportion $0.5/\sqrt{n-4}$ and $n = 10, 20, 30, 50$. With the simulation study, coverage probabilities and average length widths of confidence intervals based on robust estimator S_w^2 for the variance of nonnormal population is summarized in Tables 2-9.

In this study, Monte Carlo Simulation Method was used in obtaining the coverage probabilities and average length widths of confidence intervals due to ease of implementation for obtaining estimation values $Var(S_w^2)$. Average length widths are obtained by dividing the total of difference of the lower limit and upper limits of intervals found for each replication to the number of replications. Coverage probabilities are determined by dividing the number of cases where population variance is between the lower limit and upper limit values by the number of replications.

Two different methods were implemented for obtaining the estimator S_w^2. The first is the estimator S_w^2 which was obtained by using kurtosis coefficient G based on trimmed mean and the second is the estimator S_w^2 which was obtained by using the sample kurtosis coefficient. Coverage probabilities and average interval widths of the confidence intervals based on this estimator are summarized in the tables below.

TABLE 2: Coverage probabilities and average length widths under Student- t distribution for $\alpha = 0.05$.

	t(5)		t(10)		t(20)		t(30)	
n	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$
10	0.9535 (3.4282)	0.8487 (7.9876)	0.9577 (2.2221)	0.9109 (4.3986)	0.9538 (1.8351)	0.8671 (3.9974)	0.9518 (1.7244)	0.8589 (3.7896)
20	0.9528 (2.6488)	0.8493 (4.4587)	0.9594 (1.6546)	0.9209 (2.3681)	0.9557 (1.3610)	0.8993 (1.9749)	0.9555 (1.3097)	0.8952 (1.8926)
30	0.9518 (2.2515)	0.8515 (3.3285)	0.9554 (1.3733)	0.9339 (1.7669)	0.9541 (1.1417)	0.9152 (1.4643)	0.9523 (1.0850)	0.9107 (1.3871)
50	0.9520 (1.8030)	0.8519 (2.4289)	0.9529 (1.1031)	0.9393 (1.2982)	0.9594 (0.8999)	0.9286 (1.0473)	0.9537 (0.8618)	0.9258 (0.9973)

* Values in the parenthesis are the average of the lengths of confidence interval.

Population has Student-t distribution with parameters 5, 10, 20 and 30 when $\alpha = 0.05$. It was seen that the coverage probabilities based on kurtosis coefficient G was quite approximate to the nominal confidence level in all sample sizes. However, the coverage probabilities based on biased sample kurtosis coefficient $\hat{\gamma}$ was quite lower than the nominal confidence level in all cases. It was determined that the average length widths are reduced as the sample size increases and average length widths based on estimator S_w^2 which is obtained by using the kurtosis coefficient G are narrower (Table 2).

TABLE 3: Coverage probabilities and average length widths under Student-t distribution for $\alpha = 0.10$.

	t(5)		t(10)		t(20)		t(30)	
n	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$
10	0.9074 (2.9991)	0.8186 (6.9074)	0.9042 (1.8414)	0.8043 (4.0375)	0.8993 (1.5361)	0.7145 (3.3538)	0.9054 (1.5077)	0.8012 (3.3028)
20	0.9068 (2.3809)	0.8361 (4.2076)	0.9058 (1.3851)	0.8149 (2.0988)	0.8974 (1.1664)	0.7501 (1.6888)	0.8967 (1.1018)	0.8026 (1.5955)
30	0.9068 (1.9667)	0.8358 (3.0289)	0.9060 (1.1889)	0.8199 (1.5706)	0.8987 (0.9475)	0.7589 (1.2169)	0.8961 (0.9063)	0.8032 (1.1561)
50	0.9084 (1.4925)	0.8386 (1.9994)	0.9082 (0.9194)	0.8152 (1.0970)	0.8964 (0.7712)	0.7776 (0.8979)	0.9009 (0.7236)	0.8032 (0.8372)

* Values in the parenthesis are the average of the lengths of confidence interval.

Table 3 analyses the case where $\alpha = 0.10$. In this Table, it was determined that coverage probabilities of the confidence interval based on estimator S_w^2 which was obtained by using the kurtosis coefficient G was quite approximate to the nominal confidence level in all cases. Besides, it was also seen that the coverage probabilities increase and average length widths are reduced as the sample size increases. However, coverage probabilities based on biased sample kurtosis coefficient $\hat{\gamma}$ was quite lower than the nominal confidence level in all sample sizes (Table 3).

TABLE 4: Coverage probabilities and average length widths under Chi-squared distribution for $\alpha = 0.05$.

	χ_1^2		χ_3^2		χ_{10}^2		χ_{30}^2	
n	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$
10	0.9573 (5.3158)	0.8638 (6.0696)	0.9512 (12.5404)	0.8681 (13.6592)	0.9585 (35.1934)	0.8677 (36.7955)	0.9552 (98.9335)	0.8662 (101.2436)
20	0.9576 (4.5269)	0.8645 (4.7898)	0.9537 (9.8464)	0.8686 (10.1950)	0.9504 (26.6610)	0.8621 (27.1428)	0.9553 (73.1163)	0.8569 (73.7919)
30	0.9588 (3.8769)	0.8674 (4.0053)	0.9532 (8.2740)	0.8658 (8.4505)	0.9552 (22.7509)	0.8566 (23.0067)	0.9525 (60.5409)	0.8533 (60.8125)
50	0.9554 (3.2804)	0.8662 (3.3715)	0.9540 (6.8761)	0.8599 (6.9987)	0.9565 (18.4694)	0.8576 (18.6324)	0.9551 (48.7118)	0.8539 (48.8895)

* Values in the parenthesis are the average of the lengths of confidence interval.

In this part of the simulation study, trimming is made from only the high end of the consecutive data in obtaining the kurtosis coefficient in calculating the estimator S_w^2 since the Chi-square distribution is positively skewed distribution. It was determined in Table 4 that coverage probabilities based on kurtosis coefficient G are quite approximate to the nominal confidence level in all sample sizes for $\alpha = 0.05$. When this situation is compared with the results obtained by using the sample kurtosis coefficient $\hat{\gamma}$, it was seen that the coverage probabilities of the confidence interval based on estimator S_w^2 which was obtained by using the kurtosis coefficient G were quite approximate to the nominal confidence level and average length widths are narrower (Table 4).

TABLE 5: Coverage probabilities and average length widths under Chi-squared distribution for $\alpha = 0.10$.

	χ_1^2		χ_3^2		χ_{10}^2		χ_{30}^2	
n	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$
10	0.9001 (4.5985)	0.8545 (5.2420)	0.9002 (10.8781)	0.8526 (11.9196)	0.9044 (29.7267)	0.8266 (31.1522)	0.9078 (82.2300)	0.8146 (84.1217)
20	0.9068 (3.9293)	0.8537 (4.1597)	0.9058 (8.2205)	0.8258 (8.5143)	0.9037 (22.4265)	0.8005 (22.8434)	0.8973 (61.2001)	0.8004 (61.7616)
30	0.9069 (3.1443)	0.8377 (3.2489)	0.9052 (7.2307)	0.8163 (7.3862)	0.9083 (18.8077)	0.8017 (19.0067)	0.9007 (51.7171)	0.8016 (51.9562)
50	0.9072 (2.7033)	0.8309 (2.7803)	0.9063 (5.7745)	0.8597 (5.8793)	0.9087 (15.0437)	0.8022 (15.1734)	0.9065 (40.8599)	0.8173 (41.0062)

* Values in the parenthesis are the average of the lengths of confidence interval.

It is seen that coverage probabilities of the confidence interval based on estimator S_w^2 which was obtained by using the kurtosis coefficient G are quite approximate to the nominal confidence level even in small sample sizes. It was determined that confidence interval average length widths are reduced as the sample size increases. However, coverage probabilities based on biased sample kurtosis coefficient $\hat{\boldsymbol{\gamma}}$ was quite lower than the nominal confidence level in all sample sizes when populations had Chi-squared distribution (Table 5).

TABLE 6: Coverage probabilities and average length widths under Gamma distribution for $\alpha = 0.05$.

	Gamma (1/3,1)		**Gamma (1.5,2)**		**Gamma (2,0.5)**		**Gamma (3,1)**	
n	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$
10	0.9422 (0.7615)	0.8586 (3.9074)	0.9397 (11.9956)	0.8416 (19.8351)	0.9405 (1.3081)	0.8400 (7.4313)	0.9407 (5.9067)	0.8437 (10.5599)
20	0.9426 (0.7335)	0.8561 (2.2076)	0.9402 (10.5973)	0.8440 (17.6365)	0.9436 (1.0804)	0.8453 (6.0517)	0.9444 (4.7942)	0.8449 (6.5705)
30	0.9423 (0.3228)	0.8598 (1.0289)	0.9480 (5.5885)	0.8455 (10.4059)	0.9463 (0.6309)	0.8483 (3.5155)	0.9448 (2.6780)	0.8491 (3.7189)
50	0.9480 (0.1011)	0.8596 (1.0094)	0.9495 (2.2697)	0.8596 (5.8963)	0.9499 (0.1036)	0.8516 (2.3678)	0.9499 (1.5632)	0.8599 (2.2489)

* Values in the parenthesis are the average of the lengths of confidence interval.

Population has Gamma distribution with different parameters when $\alpha = 0.05$. It was seen that the coverage probabilities based on kurtosis coefficient G was quite approximate to the nominal confidence level in all sample sizes. However, coverage probabilities based on biased sample kurtosis coefficient $\hat{\boldsymbol{\gamma}}$ was quite lower than the nominal confidence level in all sample sizes (Table 6).

TABLE 7: Coverage probabilities and average length widths under Gamma distribution for $\alpha = 0.10$.

	Gamma (1/3,1)		Gamma (1.5,2)		Gamma (2,0.5)		Gamma (3,1)	
n	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$	***G***	$\hat{\gamma}$
10	0.8957 (0.6501)	0.8090 (3.0099)	0.9020 (6.4122)	0.8073 (12.8894)	0.9009 (0.8370)	0.8073 (3.0583)	0.9000 (4.8815)	0.8017 (6.2981)
20	0.9049 (0.6255)	0.8060 (2.8825)	0.9030 (3.9481)	0.8061 (11.5100)	0.9038 (0.6989)	0.8078 (2.8832)	0.9008 (3.9796)	0.8035 (5.9648)
30	0.9053 (0.2722)	0.8290 (1.4432)	0.9041 (2.2174)	0.8094 (6.4942)	0.9031 (0.3836)	0.8095 (1.5200)	0.9008 (2.2092)	0.8037 (3.9991)
50	0.9055 (0.1596)	0.8299 (1.1023)	0.9049 (1.8635)	0.8099 (3.2569)	0.9045 (0.1236)	0.8099 (1.1026)	0.9010 (1.9687)	0.8056 (2.1796)

In Table 7, the case where $\alpha = 0.10$ is discussed. In this table, it is determined that coverage probabilities of the confidence interval based on kurtosis coefficient G are quite close to the nominal confidence level in all sample sizes. In addition, it was seen that coverage probabilities increase and average length widths are reduced as the sample size increases (Table 7).

TABLE 8: Coverage probabilities and average length widths under Lognormal distribution for $\alpha = 0.05$.

	Lognormal(0,1)		Lognormal(0,2)		Lognormal(2,5)		Lognormal(5,4)	
n	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$
10	0.9420 (2.2253)	0.8447 (4.3551)	0.9437 (8.0066)	0.8455 (10.5949)	0.9408 (5.2110)	0.8460 (10.1285)	0.9423 (12.3620)	0.8431 (22.5986)
20	0.9493 (1.5989)	0.8482 (2.7237)	0.9453 (5.4378)	0.8481 (7.9479)	0.9462 (4.1370)	0.8485 (7.8035)	0.9440 (10.5357)	0.8481 (19.5117)
30	0.9489 (0.9548)	0.8514 (1.0355)	0.9458 (3.2649)	0.8487 (5.2432)	0.9480 (2.4260)	0.8499 (5.2786)	0.9499 (8.2147)	0.8493 (15.8540)
50	0.9500 (0.4036)	0.8515 (0.8964)	0.9510 (1.3678)	0.8516 (2.8964)	0.9510 (1.9657)	0.8510 (2.3648)	0.9520 (5.1245)	0.8510 (10.6985)

* Values in the parenthesis are the average of the lengths of confidence interval.

In terms of coverage probabilities for $\alpha = 0.05$ and Lognormal distribution, it is determined that confidence interval based on kurtosis coefficient G are quite close to the nominal confidence level in all of the sample sizes (Table 8).

TABLE 9: Coverage probabilities and average length widths under Lognormal distribution for $\alpha = 0.10$.

	Lognormal(0,1)		Lognormal(0,2)		Lognormal(2,5)		Lognormal(5,4)	
n	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$	G	$\hat{\gamma}$
10	0.9005 (1.6720)	0.8090 (3.9680)	0.9015 (6.8604)	0.8004 (12.8435)	0.9055 (42.167)	0.8010 (49.3707)	0.9057 (27.1505)	0.8092 (31.3998)
20	0.8989 (1.1402)	0.8087 (2.4174)	0.9011 (4.6125)	0.8013 (8.7524)	0.9022 (28.480)	0.8030 (35.3797)	0.9016 (18.2750)	0.8099 (22.5560)
30	0.9014 (0.7019)	0.8191 (1.8769)	0.9023 (2.7691)	0.8115 (4.5001)	0.9022 (17.398)	0.8030 (22.2619)	0.9018 (11.0300)	0.8128 (14.1234)
50	0.9020 (0.4563)	0.8110 (0.7068)	0.9040 (1.5968)	0.8130 (2.9678)	0.9030 (9.3786)	0.8030 (17.3149)	0.9020 (7.6579)	0.8130 (12.0196)

* Values in the parenthesis are the average of the lengths of confidence interval.

In Table 9, coverage probabilities and average length widths based on kurtosis coefficients G and $\hat{\gamma}$ with the random samples produced from Lognormal distribution for different sample sizes when $\alpha = 0.10$ are given. It is concluded that coverage probabilities of confidence intervals based on estimator G are quite close to the nominal confidence level in all cases. It is observed that average length widths are reduced as the sample size increases.

CONCLUSION

When confidence intervals based on robust estimator S_w^2 which was obtained using both the kurtosis coefficient based on trimmed mean and the sample kurtosis coefficient for nonnormal population variance are compared in terms of coverage probabilities, it was determined that coverage probabilities of confidence interval obtained with the estimator S_w^2 which is obtained by using the kurtosis coefficient based on trimmed mean when type I error is both $\alpha = 0.05$ and $\alpha = 0.10$ are quite approximate to the nominal confidence level even in small sample sizes where confidence interval coverage probabilities are higher. When these confidence intervals are compared in terms of average length widths, it was determined that the average length widths of the estimator S_w^2 which was obtained by using the kurtosis coefficient based on trimmed mean are narrower. According to results obtained with data produced from Chi-squared, Gamma, Lognormal and t-distributions with different distribution parameters, if it is desired to create a narrower confidence interval for nonnormal population variance, kurtosis coefficient which is obtained by using the trimmed mean should be preferred in obtaining the estimator S_w^2 which has high coverage probability. Finally it will be useful for the future studies to investigate of the coverage probabilities and average length widths of confidence intervals obtained with other distributions.

Conflict of Interest

Authors declared no conflict of interest or financial support.

Authorship Contributions

Idea/Concept: *Hayriye Esra Akyüz;* ***Design:*** *Hayriye Esra Akyüz;* ***Control/Supervision:*** *Hamza Gamgam;* ***Data Collection and/or Processing:*** *Hayriye Esra Akyüz;* ***Analysis and/or Interpretation:*** *Hayriye Esra Akyüz;* ***Literature Review:*** *Hayriye Esra Akyüz;* ***Writing the Article:*** *Hayriye Esra Akyüz;* ***Critical Review:*** *Hamza Gamgam.*

REFERENCES

1. Searls DT, Intarapanich P. A note on an estimator for the variance that utilizes the kurtosis. Amer Stat 1990;44(4):295-6.
2. Wencheko E, Chipoyera HW. Estimation of the variance when kurtosis is known. Stat Paper 2009;50(3):455-64.
3. Zou GY, Huang W, Zhang X. A note on confidence interval estimation for a linear function of binomial proportions. Comput Stat Data Anal 2009;53(4):1080-5.
4. Longford NT. Small-sample inference about variance and its transformations. Sort 2010;34(1):3-20.
5. Paksaranuwat P, Niwitpong SA. Confidence intervals for the variance and the ratio of two variances of nonnormal distributions with missing data. Thailand Stat 2010;8(1):81-92.
6. Burch BD. Confidence intervals for variance components in unbalanced one-way random effects models using non-normal distributions. J Stat Plan Inf 2011a;141(12):3793-807.
7. Burch BD. Assessing the performance of normal-based and REML-based confidence intervals for the ıntraclass correlation coefficient. Comput Stat Data Anal 2011b;55(2):1018-28.
8. Donner A, Zou GY. Closed-form confidence intervals for functions of the normal mean and standard deviation. Stat Methods Med Res 2012;21(4):347-59.
9. Niwitpong SA. A note on coverage probability of confidence interval for the difference between two normal variances. Appl Math Sci 2012;6(67):3313-20.
10. Rajić VĆ, Stanojević J. Confidence intervals for the ratio of two variances. J Appl Stat 2013;40(10):2181-7.
11. Thompson HD. Estimation of the Ratio of the Variances of Two Independent Normal Populations, MSc Thesis, University of Nevada, Reno, Published by ProQuest LLC (2014). Copyright in the Dissertation held by the Author. UMI Number: 1550932.p.1-24.
12. Suwan S, Niwitpong S. Interval estimation for a linear function of variances of nonnormal distributions that utilize the kurtosis. Appl Math Sci 2013;7(99):4909-18.
13. Banik S, Albatineh AN, Abu-Shawiesh MOA, Kibria BG. Estimating the population standard deviation with confidence interval: a simulation study under skewed and symmetric conditions. Int J Stat Med Res 2014;3(4):356-67.
14. Burch BD. Estimating kurtosis and confidence intervals for the variance under nonnormality. J Stat Comput Simul 2015;84(12):2710-20.
15. Burch BD. Distribution-dependent and distribution-free confidence intervals for the variance. Stat Meth Appl 2017;26(1):1-20.
16. Casella G, Berger RL. Statistical Inference. Duxbury Thomson Learning. 2nd ed. USA; Thomson Learning; 2001. p.686.
17. Scheffé H. The Analysis of Variance. 1st ed. New York: Wiley; 1959. p.477.
18. Royston P. Which measures of skewness and kurtosis are best? Stat Med 1992;11(3):333-43.
19. Bonett DG. Confidence interval for a ratio of variances in bivariate nonnormal distributions. J Stat Comput Simulat 2006a;76(07):637-44.
20. Bonett DG. Robust confidence interval for a ratio of standard deviations. Appl Psychol Meas 2006b;30(5):432-9.
21. Panik MJ. Advanced Statistics from an Elementary Point of View. 1st ed. Amsterdam: Elsevier/Academic Press; 2005. p.802.
22. Dalén J. Algebraic bounds on standardized sample moments. Stat Probab Lett 1987;5(5):329-31.
23. Tiku ML, Akkaya AD. Robust Estimation and Hypothesis Testing. 1st ed. New Delhi: New Age International Limited; 2004. p.354.
24. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med 2000;19(9):1141-64.
25. Efron B. Bootstrap methods: another look at the jackknife. Ann Stat 1979;7(1):1-26.