

Original Article

Improved confidence interval estimation of the population standard deviation using ranked set sampling: A simulation studyAhmed N. Albatineh^{1*}, Meredith L. Wilcox², Bashar Zogheib³, Golam B.M. Kibria⁴¹Department of Community Medicine and Behavioral Sciences, Faculty of Medicine, Kuwait University, Kuwait²Department of Epidemiology, Stempel College of Public Health, Florida International University, Miami, USA³Department of Natural Sciences, Faculty of Science, American University of Kuwait Kuwait⁴Department of Statistics, College of Science, Florida International University Miami, Florida, USA

ARTICLE INFO

Received: 22.01.2017
 Revised 12.02.2018
 Accepted 16.03.2018

Key words:

Interval estimation;
 Standard deviation;
 Ranked set sample;
 Coverage probability;
 Simulations

ABSTRACT

Background & Aim: The sample standard deviation S is the common point estimator of σ , but S is sensitive to the presence of outliers and may not be an efficient estimator of σ in skewed and leptokurtic distributions. Although S has good efficiency in platykurtic and moderately leptokurtic distributions, its classical inferential methods may perform poorly in non-normal distributions. The classical confidence interval for σ relies on the assumption of normality of the distribution. In this paper, a performance comparison of six confidence interval estimates of σ is performed under ten distributions that vary in skewness and kurtosis.

Methods and Material: A Monte Carlo simulation study is conducted under the following distributions: normal, two contaminated normal, t , Gamma, Uniform, Beta, Laplace, exponential and χ^2 with specific parameters. Confidence interval estimates obtained using the more powerful ranked set sampling (RSS) are compared with the traditional simple random sampling (SRS) technique. Performance of the confidence intervals is assessed based on width and coverage probabilities. A real data example representing birth weight of 189 newborns is used for assessment.

Results: It is not surprising that for normal data most of the intervals were close to the nominal value especially using RSS. Simulation results indicated generally better performance of RSS in terms of coverage probability and smaller interval width as sample size increases, especially for contaminated and heavy-tailed skewed distributions.

Conclusion: Simulation results revealed that the use of RSS improved greatly the coverage probability. Also, it was found that the interval labeled (III) due to Bonett (2006) had the best performance in terms of coverage probability over the wide range of distributions investigated in this paper and would be recommended for use by practitioners. There may be a need to develop nonparametric intervals that is robust against outliers and heavy-tailed distributions.

Introduction

Confidence interval estimation plays an important part in statistical inference about a population parameter and evaluating the accuracy of its estimator. A confidence interval for a parameter represents a set of plausible values for that parameter along with a confidence level. For example, if σ is the standard deviation of a population, then we seek

two endpoints L, U such that $P(L \leq \sigma \leq U) = 1 - \alpha$ for a given level α . The sample standard deviation S is the common point estimator of σ , but S is very sensitive to the presence of outliers in the data and may not be an efficient estimator of σ in skewed and leptokurtic distributions. Although S has good efficiency in platykurtic and moderately leptokurtic distributions, its classical inferential methods may perform poorly in non normal distributions, see Abou-Shawiesh et al. (1) for a discussion. The classical confidence interval for σ relies on the assumption of normality of the distribution. This interval is highly sensitive to the presence of outliers and/or

* Corresponding Author: Ahmed N. Albatineh
 Department of Community Medicine and Behavioral Sciences
 P.O Box 24923, Safat 13110, Kuwait
 Email: aalbatineh@hsc.edu.kw

departure from normality as demonstrated by Lehman and Romano (2). As a result, many researchers attempted to develop an alternative to the classical interval estimate of σ . The key question is which of those available intervals work best and under which setting? Hence there is a need for a comprehensive comparison between such interval estimates of σ .

Abou-Shawiesh et al. (1) performed a simulation study comparing seven confidence intervals for σ using data from normal, χ^2 , and Log-Normal distributions. Kittani and Zghoul (10) introduced two confidence intervals for σ based on the asymptotic distribution of the average absolute deviation from the median (AMAD) and mean absolute deviation (MAD) and compared their performance with the classical χ^2 interval over a range of distributions. Hummel et al. (9) compared the performance of four interval estimates of σ for a range of sample sizes and data generated from a number of distributions. Bonnett (5) proposed an approximate confidence interval for σ and showed that it was near exact under normality with excellent small sample properties under non normality and compared its performance with the classical χ^2 interval. Cojbasic and Tomovic (7) compared the performance of four methods for estimating confidence interval for the variance of only the exponential distribution aiming to remove most of its skewness. Citing the importance of estimating the kurtosis and its impact on the performance of the interval estimate, Burch (6) considered a number of kurtosis estimators combined with large sample theory to construct approximate confidence intervals for σ and compared performance of four interval estimates for σ over a range of distributions.

While these studies introduced new methods for estimating σ , there is a need for a more comprehensive simulation study comparing many intervals under a wide range of distributions. For that reason, six interval estimates of σ were chosen from these studies and will be compared using data generated from ten distributions in the current study, making it a more comprehensive one. Moreover, the more powerful ranked set sampling (RSS) technique will be implemented and compared with the classical simple random sampling (SRS) technique which makes this study unique in this regard. The more powerful RSS can be used to

obtain estimates of μ and σ in situations where the variable of interest is too expensive or can't be easily measured, but can be easily ranked. McIntyre (12) was the first to suggest using RSS to estimate population mean instead of SRS and the idea was later developed by Takahasi and Wakimoto (18) using mathematical theory to support their claim. Stokes (17) proposed an estimator for the variance of a ranked set sample data and showed that the estimator is asymptotically unbiased and asymptotically more efficient than the sample variance of a simple random sample of the same number of observations regardless of presence of errors in ranking. MacEachern et al. (11) proposed an alternative estimator for the variance which is unbiased and more efficient than Stokes's estimator even when the underlying distribution is not normal and the ranking of the elements is not perfect. MacEachern et al. (11) estimator of σ^2 performs well for small to moderate sample sizes and is asymptotically equivalent to Stokes (17) estimator.

In this paper, performance comparison of six confidence intervals for estimating the population σ using RSS compared with the usual SRS technique using data from ten distributions with varying skewness and kurtosis. The estimator of the variance of a ranked set sample proposed by MacEachern et al. (11) will be used in the simulations. The comparison will be based on coverage probability and intervals width which seems to be the major factors for comparison, see Albatineh et al. (3) and Terpstra and Wang (19) for a discussion.

Methods

Unbiased estimate of μ and σ using ranked set sample

Suppose that we are interested in obtaining an RSS of size n from a population. First, a SRS of size k observations are selected and rank ordered on an attribute of interest. The observation that is determined to be the smallest is the first element of the RSS and is denoted $X_{[1]1}$ and the remaining $k-1$ units are discarded. A second SRS of size k is selected from the population and ranked the same way and the second smallest observation is selected and denoted $X_{[2]1}$. In a similar fashion, $X_{[3]1}, X_{[4]1}, \dots, X_{[k]1}$ are selected, hence $X_{[1]1}, X_{[2]1}, \dots, X_{[k]1}$ represent our first balanced RSS of set size k . To obtain a balanced RSS of size $n=km$, the process is repeated m

Table 1. Balanced RSS with m cycles and set size k

Cycle 1	X _{[1]1}	X _{[2]1}	X _{[3]1}	...	X _{[k]1}
Cycle 2	X _{[1]2}	X _{[2]2}	X _{[3]2}		X _{[k]2}
Cycle 3	X _{[1]3}	X _{[2]3}	X _{[3]3}		X _{[k]3}
⋮	⋮	⋮			⋮
⋮	⋮	⋮			⋮
Cycle m	X _{[1]m}	X _{[2]m}	X _{[3]m}		X _{[k]m}

independent cycles yielding the balanced RSS of size n shown in Table 1.

The complete balanced RSS with set size k and m cycles is given by $\{X_{[r]i}: r=1, 2, \dots, k; i=1, 2, \dots, m\}$. The term $X_{[r]i}$ is called the r -th judgment order statistic from the i -th cycle. It is the observation that is judged to be the r -th order statistic from one of the k sets in the i -th cycle, see MacEachern et al. (11) for discussion. Assume that the underlying distribution has finite mean μ and variance σ^2 , Stokes (17) proposed an estimator of σ^2 based on RSS given by

$$\hat{\sigma}^2 = \frac{1}{km-1} \sum_{i=1}^m \sum_{r=1}^k (X_{[r]i} - \hat{\mu})^2, \text{ where } \hat{\mu} = \frac{1}{km} \sum_{i=1}^m \sum_{r=1}^k X_{[r]i} \quad (1)$$

Stokes (17) showed that this estimator is a biased estimator of σ^2 , but it is asymptotically unbiased as either k or m approach ∞ . Moreover, Stokes (17) indicated that the RSS estimator $\hat{\mu}$ has more precision over the sample mean; say \bar{y} obtained using SRS because of independence of the order statistics composing the ranked set sample. In fact, the author showed that $\text{var}(\bar{y}) \geq \text{var}(\hat{\mu})$. The balanced RSS is used in this paper. The estimator of the variance of a RSS proposed by MacEachern et al. (11) will be implemented in the simulations since it has been shown to perform very well for small as well as large ranked set samples, which is an improvement on Stokes estimator (17). This estimator is given by

$$\hat{\sigma}^2 = \frac{1}{km} ((k-1)MST + (mk - k + 1)MSE) \quad (2)$$

where MST (Mean Square Treatment) and MSE (Mean Square Error) are obtained from an analysis of variance performed on the ranked set sample data with the judgment class used as the factor and are given by

$$MST = \frac{1}{k-1} \sum_{i=1}^m \sum_{r=1}^k (X_{[r]i} - \hat{\mu})^2 - \frac{1}{k-1} \sum_{i=1}^m \sum_{r=1}^k (X_{[r]i} - \bar{X}_{[r]})^2 \quad (3)$$

$$MSE = \frac{1}{k(m-1)} \sum_{r=1}^k \sum_{i=1}^m (X_{[r]i} - \bar{X}_{[r]})^2, \text{ where } \bar{X}_{[r]} = \frac{1}{m} \sum_{i=1}^m X_{[r]i} \quad (4)$$

Confidence intervals for σ

In this section, six widely used confidence intervals for estimating σ are presented.

Exact confidence interval (I)

Let $x_1, x_2, x_3, \dots, x_n$ be an independent and identically distributed random sample from a normal distribution with finite mean μ and variance σ^2 , i.e. $X \sim N(\mu, \sigma^2)$, then

$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2 \sim \chi^2(n-1) \quad (5)$$

Therefore,

$$P(\chi^2_{(\frac{\alpha}{2})} \leq \chi^2 \leq \chi^2_{(1-\frac{\alpha}{2})}) = P(\chi^2_{(\frac{\alpha}{2})} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{(1-\frac{\alpha}{2})}) \quad (6)$$

where $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ is the sample variance, Hence, a $(1-\alpha)100\%$ exact confidence interval for the population variance is given by

$$P\left(\frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2}, n-1)}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2}, n-1)}}\right) = 1 - \alpha \quad (7)$$

where $\chi^2_{(\frac{\alpha}{2})}$ and $\chi^2_{(1-\frac{\alpha}{2})}$ are the $(\frac{\alpha}{2})$ th and $(1-\frac{\alpha}{2})$ th percentile of the χ^2 distribution with $n - 1$ degrees of freedom. Taking the square root of equation (7) gives the $(1-\alpha)100\%$ confidence interval for σ , marked (I) in the tabulated simulations, which is given by

$$\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2}, n-1)}}} < \sigma < \sqrt{\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2}, n-1)}}} \right) \quad (8)$$

Robust confidence interval (II)

Let $x_1, x_2, x_3, \dots, x_n$ be an independent and identically distributed random sample from a distribution function F . For such sample the Q_n is defined as

$$Q_n = 2.2219 \{ |x_i - x_j| ; i < j ; i = 1, 2, \dots, n ; j = 1, 2, \dots, n \}_g \tag{9}$$

where $g = \binom{h}{2} \binom{n}{2} / 4$, and $h = \left[\frac{h}{2} \right] + 1$ and the symbols $(.)$ represents combination, and $[.]$ is the integer value. The Q_n estimator is the g^{th} order statistic of the $\binom{n}{2}$ integer point distances. The

value 2.2219 was chosen to make Q_n a consistent estimator of scale for normal data. Rousseeuw and Croux (14) derived the factor d_n which makes the quantity $d_n Q_n$ unbiased estimator of σ for the case of normal distribution. The authors provided approximate values of d_n for larger values of n which is given by

$$d_n = \left\{ \begin{array}{l} \frac{n}{n+1.4}, \text{ if } n \text{ is odd} \\ \frac{n}{n+3.8}, \text{ if } n \text{ is even} \end{array} \right\} \tag{10}$$

Let $x_1, x_2, x_3, \dots, x_n$ be an independent and identically distributed random sample of size n from a continuous distribution. Define the random variable T as

$$T = \frac{d_n Q_n}{\sigma} \tag{11}$$

where $d_n Q_n$ is the unbiased estimator of σ so that $E(T) = 1$ for normal distribution. Rousseeuw and Croux (14) showed that for large n , the following asymptotic result holds:

$$T = \frac{d_n Q_n}{\sigma} \square N \left(1, \frac{1}{1.65n} \right) \Rightarrow d_n Q_n \square N \left(\sigma, \frac{1}{1.65n} \sigma^2 \right) \tag{12}$$

Following the development in Abu-Shawiesh et al. (1), a $(1-\alpha)100\%$ robust confidence interval for σ , marked (II) in the tabulated simulations, is given by

$$\left(\frac{1.28\sqrt{n} * d_n Q_n}{Z_{1-\frac{\alpha}{2}} + 1.28\sqrt{n}}, \frac{1.28\sqrt{n} * d_n Q_n}{Z_{\frac{\alpha}{2}} + 1.28\sqrt{n}} \right) \tag{13}$$

where $Z_{1-\frac{\alpha}{2}}$ and $Z_{\frac{\alpha}{2}}$ are the $(1-\frac{\alpha}{2})$ and $(\frac{\alpha}{2})$ th percentiles of the standard normal distribution.

Bonett confidence interval (III)

Let $x_1, x_2, x_3, \dots, x_n$ be an independent and identically distributed random sample from a normal distribution, i.e. $x_i \square N(\mu, \sigma^2)$. Mood et al. (13) indicated that the variance of S^2 can be expressed as $\sigma^4 \{ \gamma_4 - (n-3) / (n-1) \} / n$, where $\gamma_4 = \mu_4 / \sigma^4$ where μ^4 is the population fourth central moment. A variance-stabilizing transformation for S^2 is $\ln(S^2)$ and using the δ -method gives $\text{var}(\ln(S^2)) \cong \{ \gamma_4 - (n-3) / (n-1) \} / n$. Shoemaker (16) found out that using the quantity $\{ \gamma_4 - (n-3) / n \} / (n-1)$ improved the small sample performance of his equal-variance test, and this small-sample adjustment will be used in our simulations. To estimate $\text{var}(\ln S^2)$, one needs to estimate γ_4 since it is unknown in practice.

The Pearson's estimator $\hat{\gamma}_4 = (n \sum_{i=1}^n (x_i - \bar{x})^4) / (\sum_{i=1}^n (x_i - \bar{x})^2)^2$ tends to have large negative bias in leptokurtic (heavy tailed) distributions unless we have a large sample size. For that reason, Bonett (5) proposed an estimator for γ_4 which is asymptotically equivalent to Pearson's estimator and is given by

$$\hat{\gamma}_4 = \frac{n \sum_{i=1}^n (x_i - m)^4}{(\sum_{i=1}^n (x_i - \bar{x})^2)^2} \tag{14}$$

where m is the trimmed mean with trim-proportion equals $1 / (2\sqrt{n-4})$. Bonett (5) indicated that this estimator of kurtosis tends to have less negative bias and smaller coefficient of variability compared to Pearson's estimator in symmetric and skewed leptokurtic distributions, and will be implemented in our simulations. Bonett (5) proposed the confidence interval

limits for σ^2 given by

$$LCL = \exp\{\ln cS^2 - Z_{\alpha/2}SE\} \text{ and } UCL = \exp\{\ln cS^2 + Z_{\alpha/2}SE\} \quad (15)$$

where $Z_{\alpha/2}$ is the two-sided critical value from the standard normal distribution, the standard error

$$SE = \sqrt{\hat{\gamma}_4 - (n-3)/n} / (n-1), \text{ where } c = n / (n - Z_{\alpha/2})$$

and $\hat{\gamma}_4$ as given in equation (14). Taking the square root of the endpoints of equation (15) gives the endpoints of (1- α)100% confidence interval for σ which will be marked (III) in the tabulated simulations.

Large sample normal approximation (IV)

It is well known that the distribution of the sample variance S^2 is asymptotically normally distributed with expected value σ^2 and variance $(\gamma-1)\sigma^4/n$, provided that the fourth moment of the parent distribution is finite and γ is the population kurtosis, see Hummel et. al. (9) and Arnold (4) for a discussion. Such an asymptotic distribution gives rise to the (1- α)100% confidence interval for the variance given by

$$\left(\frac{S^2}{1 + Z_{1-\alpha/2}\sqrt{\frac{\hat{\gamma}-1}{n}}} \leq \sigma^2 \leq \frac{S^2}{1 - Z_{1-\alpha/2}\sqrt{\frac{\hat{\gamma}-1}{n}}} \right) \quad (16)$$

where $\hat{\gamma}$ is a consistent estimator of the kurtosis which is unbiased for normal samples and $\hat{\gamma}_e$ is the corresponding estimate of the excess kurtosis ($\gamma=3$ and $\gamma_e=0$ for normal distribution) given by

$$\hat{\gamma}_e = \hat{\gamma} - 3 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \frac{(x_i - \bar{x})^4}{S^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

Taking the square root of equation (16) gives the (1- α)100% confidence interval for σ which will be marked "IV" in the tabulated simulations.

Adjusting for skewness (V)

Note that for small samples, the distribution of S^2 has high skewness. To help adjust for this skewness, the natural logarithm transformation is applied to the sample variance. Using the Cramer δ method, $\ln(S^2)$ will be asymptotically approximately normally distributed with mean $\ln(\sigma^2)$ and corresponding variance $(\gamma-1)/n$. Using this transformation, a (1- α)100% confidence interval for σ^2 is given by

$$S^2 \exp\left(-Z_{1-\alpha/2}\sqrt{\frac{\hat{\gamma}-1}{n}}\right) \leq \sigma^2 \leq S^2 \exp\left(Z_{1-\alpha/2}\sqrt{\frac{\hat{\gamma}-1}{n}}\right) \quad (17)$$

Taking the square root of equation (17) gives the (1- α)100% confidence interval for σ which will be marked (V) in the tabulated simulations.

Adjusted degrees of freedom (VI)

It is known that the sample variance is a sum of squares, and for sufficiently large samples, it can be approximated by a χ^2 distribution with an appropriate degrees of freedom which can be estimated using the matching of moments method. A similar argument is used by Shoemaker (16) when approximating the distribution of a ratio of independent sample variances. The first two moments of the distribution of S^2 are matched with those of a random variable $Y \sim c\chi_r^2$, and then solving for r and c the equations $\sigma^2 = cr$, and $\frac{\sigma^4}{n}(\gamma - \frac{n-3}{n-1}) = 2rc^2$. Mood et. al. (13) indicated that when sampling from any distribution with finite first four moments, we have

$$\text{var}(S^2) = \frac{\sigma^4}{n} \left(\gamma - \frac{n-3}{n-1}\right) \quad (18)$$

The unique solution is given by

$$r = \frac{2n(n-1)}{(n-1)\gamma - n + 3} = \frac{2n(n-1)}{(n-1)\gamma_e + 2n}$$

$$c = \frac{\sigma^2}{2n} \left(\gamma - \frac{n-3}{n-1}\right) = \frac{\sigma^2}{2n} \left(\gamma_e + \frac{2n}{n-1}\right) \quad (19)$$

Therefore, $\frac{rS^2}{\sigma^2}$ is approximately distributed as χ_r^2 , and hence an approximate confidence interval estimate for the variance is given by

$$\frac{\hat{r}S^2}{\chi_{\hat{r}, 1-\alpha/2}^2} \leq \sigma^2 \leq \frac{\hat{r}S^2}{\chi_{\hat{r}, \alpha/2}^2}, \text{ where } \hat{r} = \frac{2n}{\hat{\gamma}_e + 2n / (n-1)} \quad (20)$$

Taking the square root of equation (20) gives the (1- α)100% confidence interval for σ which will be marked (VI) in the tabulated simulations

Results

Confidence interval estimation is one form of estimation in which an interval of plausible values of the population parameter of interest is

given, attached to a confidence level. The most common criteria for comparing confidence intervals are coverage probabilities and width. In simulation studies, it is desired that the coverage probability be very close to the nominal confidence level. When the coverage probability is greater than the nominal confidence level, the confidence interval is a conservative one and is considered to be valid, but it leads to confidence intervals that are wider than they are supposed to be. A confidence interval is anti-conservative if it is associated with a coverage probability that is smaller than the nominal confidence level. Such a confidence interval is not valid, and it generally produces confidence intervals shorter than they need to be.

Simulation Study

It is hard to perform a theoretical comparison between many confidence intervals; hence a simulation study will be conducted for such purpose. A range of random sample sizes: $n=20, 30, 50, 100$ will be generated from the following distributions:

1. $N(0, 1)$: Normal distribution with mean zero and standard deviation one.
2. $t(10)$: t distribution with 10 degrees of freedom.
3. $Beta(3,3)$: Beta distribution with parameters shape=3 and scale =3.
4. $Laplace(0,1)$: Laplace distribution (double exponential) with location parameter $\mu=0$ and scale parameter $b=1$.
5. $CN(0,0.95)$: Contaminated normal distribution
 $CN(\mu, 1-\xi) = (1-\xi)N(\mu, \sigma_1) + \xi N(\mu, \sigma_2)$
with $\mu=0, \sigma_1=1, \sigma_2=2$ and contamination proportion $\xi=5\%$.
6. $CN(0,0.90)$: Contaminated normal distribution
 $CN(\mu, 1-\xi) = (1-\xi)N(\mu, \sigma_1) + \xi N(\mu, \sigma_2)$
with $\mu=0, \sigma_1=1, \sigma_2=2$ and contamination proportion $\xi=10\%$.
7. $Exp(1)$: Exponential distribution with rate parameter $\lambda=1$.
8. $\chi^2(3)$: χ^2 distribution with 3 degrees of freedom.
9. $U(1,5)$: Uniform distribution with minimum=1 and maximum=5.
10. $Gamma(2,2)$: Gamma distribution with shape parameter $k=2$ and scale parameter $\theta=2$.

Using 2000 replications, the simulation error

for a two-sided 95% confidence interval is

$$\sqrt{\frac{(1-0.95)*0.95}{2000}} = 0.00487 \approx 0.49\%$$

. Thus, any coverage probability between 0.9404 and 0.9596 will be within a set of acceptable values. Of course a coverage probability higher than 0.9596 will still imply a valid interval but it simply means that such interval is wider than it is supposed to be. An interval with coverage smaller than 0.9404 imply that such interval is not valid and is shorter than what it is supposed to be. Given two confidence intervals for σ with roughly the same coverage probability, only then we will resort to the interval width to choose which interval is better.

Ten distributions were considered in the simulations which represents: symmetric distributions {Normal (0,1)}, heavy-tailed symmetric distributions {t(10), Laplace (0,1), CN(0,1,0.95), CN(0,2,0.90)}, light-tailed symmetric distributions {Beta (3,3), Uniform (1,5)}, light-tailed skewed distributions {Exponential (1)}, and heavy tailed skewed

distributions { $\chi^2(3)$, Gamma (2,2)}. Below is a summary for the findings for each group of distributions.

Symmetric Distributions

Since interval (I) is derived under the assumption of normality, it is not surprising to see its coverage probability close to the nominal value using both RSS and SRS for both small and large sample sizes. Interval (II) achieved nominal value coverage for sample size ≥ 50 , using both RSS and SRS, while intervals (III), (IV), (V), (VI) achieved coverage probability closer to nominal value using RSS for sample sizes ≤ 50 and using both RSS and SRS for $n > 50$ except interval (III) using SRS.

Heavy-Tailed Symmetric Distributions

For the t-distribution, coverage obtained using RSS were closer to the 95% nominal target for intervals (I), (III), (IV), and (VI) for $n=20$, but for $n \geq 50$, all intervals except (I) attained nominal coverage probability. For Laplace distribution, coverage using RSS were still higher than SRS, but were closer to nominal value when $n \geq 50$ using intervals (III), (IV), and (VI). For contaminated normal distributions, coverage using RSS is still higher than SRS, but got worse for 10% contamination especially for

Table 2. Estimated coverage probabilities and intervals widths of six confidence intervals for σ with data from Normal(0,1), t(10), Beta (3,3), Laplace (0,1) for sample sizes 20,30,50, and 100 using RSS and SRS techniques. The row with (*) represents coverage probability and the row with (**) represents interval width.

Normal (0,1)							t(10)					
Interval	I	II	III	IV	V	VI	I	II	III	IV	V	VI
n=20							n=20					
RSS	0.967 ^Y	0.928	0.978	0.966	0.948	0.959	0.943	0.932	0.969	0.956	0.912	0.947
SRS	0.681 ^Z	0.796	0.779	1.158	0.656	0.744	0.744	0.853	0.882	1.689	0.757	0.867
	0.950 ^Y	0.925	0.696	0.916	0.901	0.918	0.927	0.912	0.535	0.922	0.875	0.911
n=30							n=30					
RSS	0.957	0.936	0.981	0.943	0.935	0.942	0.917	0.922	0.963	0.936	0.922	0.930
SRS	0.543	0.613	0.637	0.628	0.508	0.547	0.606	0.659	0.765	0.898	0.634	0.693
	0.540	0.600	0.244	0.617	0.498	0.538	0.906	0.907	0.470	0.922	0.900	0.916
n=50							n=50					
RSS	0.977	0.970	0.991	0.976	0.964	0.972	0.926	0.949	0.979	0.965	0.947	0.957
SRS	0.404	0.459	0.483	0.459	0.407	0.426	0.450	0.488	0.592	0.632	0.521	0.552
	0.408	0.453	0.177	0.431	0.388	0.405	0.913	0.915	0.429	0.938	0.922	0.933
n=100							n=100					
RSS	0.962	0.957	0.985	0.959	0.955	0.959	0.914	0.920	0.974	0.950	0.944	0.951
SRS	0.282	0.315	0.341	0.294	0.280	0.286	0.315	0.336	0.428	0.398	0.368	0.378
	0.283	0.314	0.118	0.289	0.276	0.281	0.893	0.891	0.385	0.929	0.929	0.929
Beta (3,3)							Laplace (0,1)					
Interval	I	II	III	IV	V	VI	I	II	III	IV	V	VI
n=20							n=20					
RSS	0.990	0.957	0.991	0.975	0.956	0.973	0.836	0.832	0.910	0.933	0.848	0.897
SRS	0.128	0.154	0.136	0.149	0.108	0.121	0.909	0.947	1.168	2.569	1.033	1.204
	0.131	0.152	29.122	0.126	0.098	0.109	0.825	0.794	0.252	0.883	0.829	0.862
n=30							n=30					
RSS	0.983	0.951	0.984	0.947	0.947	0.955	0.832	0.757	0.931	0.912	0.888	0.905
SRS	0.102	0.120	0.109	0.092	0.081	0.086	0.755	0.724	1.080	1.694	0.947	1.061
	0.102	0.118	4.634	0.090	0.079	0.085	0.791	0.712	0.210	0.898	0.859	0.889
n=50							n=50					
RSS	0.997	0.965	0.999	0.984	0.979	0.984	0.844	0.631	0.955	0.959	0.924	0.941
SRS	0.076	0.090	0.081	0.069	0.064	0.067	0.563	0.531	0.867	1.158	0.801	0.868
	0.077	0.089	2.055	0.065	0.061	0.063	0.793	0.560	0.179	0.913	0.892	0.900
n=100							n=100					
RSS	0.992	0.957	0.995	0.978	0.972	0.975	0.818	0.362	0.964	0.944	0.931	0.936
SRS	0.053	0.062	0.057	0.045	0.043	0.044	0.398	0.363	0.651	0.683	0.592	0.617
	0.053	0.061	1.007	0.044	0.043	0.043	0.796	0.367	0.158	0.926	0.911	0.917

larger n . For 5% contamination, intervals (I), (III), (V), and (VI) did well but only interval (III) remained closer to nominal value as n increased. For 10% contamination, interval (III) has the best coverage but still lower than nominal value.

Light-Tailed Symmetric Distributions

For Beta distribution, intervals (I) and (III) always produced wider intervals than supposed

to be using both RSS and SRS for all sample sizes considered. Interval (II) produced coverage very close to nominal value using both RSS and SRS for all sample sizes. Intervals (IV), (V), and (VI), although had valid coverage but their performances were mixed, i.e. closeness to nominal value depending on the sample size. For the Uniform distribution, intervals (I), (II), and (III) produced coverage probability higher than the nominal value, though valid but wider than

Table 3. Estimated coverage probabilities and width of six confidence intervals for σ with data from contaminated normal, exponential ($\lambda = 1$), and $\chi^2(3)$ distributions for sample sizes 20,30,50, and 100 using RSS and SRS techniques. The row with (*) represents coverage probability and the row with (**) represents interval width.

Interval	Contaminated Normal CN(0,1,0.95)						Contaminated Normal CN(0,1,0.90)					
	I	II	III	IV	V	VI	I	II	III	IV	V	VI
	n=20						n=20					
RSS	0.943*	0.906	0.969	0.934	0.941	0.946	0.883	0.858	0.936	0.877	0.919	0.904
	0.717**	0.825	0.838	1.432	0.711	0.812	0.752	0.859	0.895	1.524	0.766	0.877
SRS	0.902*	0.884	0.565	0.863	0.893	0.884	0.840	0.863	0.476	0.814	0.881	0.856
	0.733**	0.814	0.303	1.128	0.667	0.754	0.769	0.847	0.287	1.285	0.721	0.818
	n=30						n=30					
RSS	0.876	0.900	0.964	0.882	0.929	0.914	0.784	0.848	0.899	0.796	0.879	0.843
	0.580	0.635	0.727	0.914	0.600	0.657	0.614	0.662	0.791	1.032	0.662	0.729
SRS	0.870	0.896	0.489	0.874	0.907	0.892	0.772	0.835	0.376	0.763	0.856	0.814
	0.578	0.630	0.223	0.837	0.578	0.631	0.614	0.660	0.206	0.919	0.641	0.702
	n=50						n=50					
RSS	0.891	0.929	0.977	0.920	0.958	0.939	0.752	0.866	0.928	0.818	0.914	0.864
	0.431	0.476	0.558	0.603	0.487	0.516	0.456	0.493	0.623	0.731	0.554	0.592
SRS	0.851	0.901	0.428	0.867	0.915	0.896	0.677	0.807	0.274	0.707	0.804	0.752
	0.437	0.473	0.158	0.559	0.467	0.493	0.467	0.494	0.143	0.666	0.535	0.568
	n=100						n=100					
RSS	0.791	0.912	0.945	0.848	0.908	0.874	0.554	0.818	0.810	0.652	0.763	0.702
	0.302	0.327	0.405	0.373	0.346	0.355	0.320	0.339	0.453	0.433	0.395	0.407
SRS	0.776	0.881	0.354	0.812	0.878	0.851	0.530	0.788	0.173	0.598	0.709	0.652
	0.302	0.325	0.103	0.367	0.340	0.349	0.322	0.339	0.093	0.426	0.390	0.402
	Exponential ($\lambda = 1$)						$\chi^2(3)$					
Interval	I	II	III	IV	V	VI	I	II	III	IV	V	VI
	n=20						n=20					
RSS	0.767	0.563	0.839	0.874	0.746	0.816	0.837	0.821	0.884	0.888	0.795	0.854
	0.607	0.536	0.767	1.684	0.666	0.774	1.548	1.514	1.902	4.189	1.641	1.901
SRS	0.735	0.523	0.513	0.815	0.720	0.768	0.810	0.749	0.084	0.834	0.766	0.801
	0.627	0.534	0.350	1.623	0.656	0.759	1.578	1.494	0.142	3.559	1.568	1.800
	n=30						n=30					
RSS	0.749	0.294	0.885	0.887	0.821	0.855	0.817	0.653	0.912	0.883	0.850	0.870
	0.511	0.401	0.750	1.376	0.660	0.746	1.296	1.153	1.834	3.183	1.594	1.790
SRS	0.738	0.329	0.366	0.866	0.815	0.838	0.796	0.609	0.078	0.862	0.810	0.839
	0.516	0.403	0.224	1.391	0.661	0.747	1.273	1.141	0.095	2.860	1.508	1.684
	n=50						n=50					
RSS	0.745	0.05	0.912	0.917	0.874	0.896	0.808	0.396	0.930	0.923	0.891	0.906
	0.392	0.295	0.659	1.204	0.620	0.686	0.979	0.850	1.546	2.230	1.436	1.570
SRS	0.722	0.124	0.297	0.884	0.856	0.873	0.765	0.398	0.053	0.883	0.862	0.876
	0.399	0.297	0.142	0.998	0.603	0.662	0.983	0.846	0.061	1.966	1.348	1.462
	n=100						n=100					
RSS	0.706	0.001	0.931	0.913	0.901	0.909	0.788	0.067	0.951	0.933	0.915	0.921
	0.281	0.202	0.526	0.635	0.491	0.521	0.687	0.581	1.166	1.270	1.067	1.118
SRS	0.704	0.008	0.253	0.913	0.897	0.912	0.747	0.116	0.056	0.910	0.899	0.905
	0.281	0.201	0.083	0.632	0.488	0.518	0.690	0.580	0.037	1.275	1.060	1.111

they should be. Intervals (IV), (V), and (VI) had coverage closer to nominal value with RSS performing better than SRS, especially for $n < 50$.

Light-Tailed Skewed Distributions

For the exponential distribution, coverage was smaller than the nominal value for all intervals, with best performance by intervals (III) and (IV) especially for $n \geq 50$. The RSS produced coverage probability higher than SRS, although lower than the nominal value.

Heavy Tailed Skewed Distributions

For the χ^2 distribution, the coverage probability of intervals (III), (IV), (V), and (VI) increased with an increase in sample size, while coverage of (I) and (II) decreased with an increase in sample size. RSS produced coverage probability higher than that of SRS. Interval (III) had the best performance especially for $n = 100$. The pattern in the Gamma distribution is almost identical to that of the χ^2 distribution with interval (III) as best performer followed by

Table 4. Estimated coverage probabilities and width of six confidence intervals for σ with data from Uniform (1,5) and Gamma (2,2) distributions for sample sizes 20,30,50, and 100 using RSS and SRS techniques. The row with (*) represents coverage probability and the row with (**) represents interval width.

Interval	Uniform (1,5)						Gamma (2,2)					
	I	II	III	IV	V	VI	I	II	III	IV	V	VI
	n=20						n=20					
RSS	0:997*	0.986	0.994	0.940	0.952	0.960	0.880	0.895	0.921	0.910	0.852	0.886
	0:781**	0.909	0.744	0.551	0.483	0.530	1.834	1.877	2.233	4.398	1.924	2.222
SRS	0:995*	0.988	0.738	0.931	0.928	0.939	0.839	0.822	0.065	0.837	0.784	0.817
	0:780**	0.883	0.309	0.528	0.466	0.513	1.866	1.841	0.120	3.805	1.819	2.082
	n=30						n=30					
RSS	0.997	0.995	0.997	0.955	0.953	0.960	0.851	0.818	0.934	0.900	0.860	0.887
	0.630	0.725	0.600	0.421	0.390	0.415	1.497	1.426	2.043	3.103	1.750	1.952
SRS	0.998	0.995	0.720	0.924	0.928	0.940	0.832	0.772	0.056	0.860	0.841	0.855
	0.630	0.715	0.249	0.413	0.383	0.409	1.502	1.427	0.082	3.061	1.716	1.908
	n=50						n=50					
RSS	0.999	0.999	0.998	0.983	0.977	0.983	0.845	0.664	0.944	0.930	0.900	0.917
	0.467	0.546	0.444	0.328	0.312	0.324	1.127	1.055	1.682	2.303	1.539	1.666
SRS	0.998	0.997	0.770	0.949	0.945	0.954	0.816	0.606	0.051	0.902	0.887	0.899
	0.472	0.536	0.187	0.305	0.293	0.304	1.149	1.051	0.053	2.285	1.524	1.649
	n=100						n=100					
RSS	1.000	0.999	1.000	0.973	0.970	0.976	0.826	0.304	0.962	0.937	0.925	0.934
	0.326	0.374	0.309	0.212	0.208	0.212	0.795	0.726	1.270	1.334	1.146	1.195
SRS	0.997	0.998	0.752	0.939	0.944	0.947	0.801	0.310	0.043	0.921	0.899	0.908
	0.327	0.371	0.129	0.208	0.204	0.208	0.794	0.721	0.034	1.299	1.127	1.174

intervals (IV) and (VI), respectively.

Real Data Example

The birth weight data used in this example is obtained from Hosmer and Lemeshow (8), which was collected from the Baystate Medical Center in Springfield, Massachusetts (University

of Massachusetts Amherst). In this data, a baby weighing less than 2500 grams is defined as "low birth weight" child. Data were collected from 189 women of which 59 women had low birth weight babies and 130 women had normal birth weight babies. For this data, the average birth weight was 2944.66 grams, with a standard

Table 5. Tests of normality results for the birth weight data

Test of normality	Test statistic	P value
Shapiro - Wilk	W = 0.9925	0.4384
Anderson - Darling	A = 0.4157	0.3301
Cramer - von Mises	W = 0.0584	0.395
Kolmogorov - Smirnov	D = 0.0435	0.5169

Table 6. Estimated six confidence intervals for σ along with their widths using birth weight data with sample sizes 25, 50, and 100 selected using RSS and SRS techniques.

Interval	I	II	III	IV	V	VI
RSS	n=25					
Width	(480.7,856.5)	(547.6,1031.1)	(490.8,838.0)	(513.4,821.4)	(494.5,766.5)	(502.5,795.0)
	375.7	483.5	347.2	308.0	271.9	292.4
SRS	(531.6,947.2)	(592.7,1116.0)	(709.2,709.2)	(567.1,910.9)	(546.0,848.9)	(555.1,880.8)
Width	415.5	523.3	0.001	343.7	302.9	325.7
RSS	n=50					
Width	(525.4,783.8)	(577.1,896.1)	(508.7,809.4)	(531.0,814.0)	(514.2,769.3)	(523.1,789.0)
	258.4	319.0	300.6	282.9	255.1	265.9
SRS	(617.5,921.3)	(624.6,969.9)	(754.2,754.2)	(619.4,974.6)	(597.9,914.1)	(609.4,940.0)
Width	303.7	345.3	0.0002	355.1	316.2	330.5
RSS	n=100					
Width	(592.8,784.3)	(616.2,839.1)	(587.5,791.3)	(613.4,760.3)	(607.4,750.4)	(610.2,755.7)
	191.5	222.8	203.7	146.8	142.9	145.4
SRS	(634.7,839.8)	(620.0,844.2)	(730.1,730.1)	(628.6,878.3)	(615.3,849.4)	(622.5,862.3)
Width	205.0	224.2	0.0001	249.6	234.1	239.8

deviation of 729.02 grams. Four tests of normality, as shown in Table 5, indicated that the birth weight data follow a normal distribution.

Table 6 presents six confidence intervals of σ for the birth weight data with samples of size 25, 50, and 100 using RSS and SRS. All intervals captured, $S=729.02$, the point estimate of σ , except for the third interval using SRS. Also, there is a clear pattern of smaller width of the interval as the sample size increase from 25 to 100. Also, the intervals produced by RSS have smaller width compared to that of SRS.

Discussion

Parameters' estimation using confidence intervals is more desirable than merely testing if the parameter equals a specified value. Interval estimation gives the practitioners an idea about the set of plausible values of the parameter with some confidence. Several papers discussed estimating σ using confidence interval approach, but few of these studies were comprehensive in the sense of conducting a comprehensive simulation study that involves several distributions, see Abu-Shawiesh et al. (1), Bonett (5), and Hummel et al. (9) to name a few. Moreover, none of those studies implemented the more powerful RSS in estimating σ . Several types of distributions including symmetric, heavy-tailed symmetric, light-tailed symmetric, light-tailed and heavy tailed skewed were included. Simulations results revealed that the type of distribution had some role in reaching nominal coverage along with the sample size considered. In general, for non normal distributions the coverage probability improved with larger sample size but with RSS intervals achieved better coverage probability and that is closer to the nominal value. When the data was contaminated at 10%, the coverage probability was affected with interval III having the best coverage but still lower than nominal value. This may indicate the need to look for some nonparametric interval estimators for σ especially for small n and presence of contamination or presence of light or heavy skewness.

Conclusion

In this paper, a simulation study was conducted to compare six well known confidence interval estimates for the population

standard deviation. Data were generated from ten distributions with varying skewness and kurtosis and sample sizes 20, 30, 50, and 100 to capture a wider picture of the performance. A comparison of ranked set sampling against simple random sampling in estimating six confidence intervals for the population standard deviation is performed. Two main criteria for comparison were implemented, namely: coverage probability and confidence interval width. Simulation results revealed that the use of RSS improved greatly the coverage probability. Also, it was found that the interval labeled (III) due to Bonett (5) had the best performance in terms of coverage probability over the wide range of distributions investigated in this paper and would be recommended for use by practitioners.

Conflict of Interests

The authors have no conflict of interests.

Acknowledgment

The code for generating contaminated normal data was provided by Professor J.C. Wang available at:

<http://www.stat.wmich.edu/wang/R/codes/index.html>.

References

1. Abu-Shawiesh MOA, Banik S, Kibria BMG. A simulation study on some confidence intervals for the population standard deviation. SORT 2011; 35(2): 83 -102.
2. Lehmann EL, Romano JP. Testing statistical hypotheses. Springer Science & Business Media; 2006.
3. Albatineh AN, Kibria BMG, Wilcox ML, Zogheib B. Confidence interval estimation for the population coefficient of variation using ranked set sampling: a simulation study. J Appl Stat. 2013; 41(4): 733-751.
4. Arnold SF. Mathematical Statistics. Prentice Hall College Div; 1990.
5. Bonett DG. Approximate confidence interval for standard deviation of non-normal distributions. Comput Stat Data Anal.2006; 50(3): 775-782.
6. Burch BD. Estimating kurtosis and confidence intervals for the variance under non-normality. J Stat Comput Simul.2013; 84(12): 2710-2720.
7. Cojbasic V, Tomovic A. Nonparametric confidence intervals for population variance of one sample and the difference of variances of two samples. J Comput Stat Data Anal.2007; 51(12): 5562-5578.

8. Hosmer D W, Lemeshow S. Applied Logistic Regression: Second edition. New York: John Wiley & Sons, 2000.
9. Hummel R, Banga S, Hettmansperger TP. Better confidence intervals for the variance in a random sample. Minitab Technical Report. (2005) Retrieved from:
<http://www.minitab.com/support/documentation/answers/OneVariance.pdf>
10. Kittani H, Zghoul A. A robust confidence interval for the population standard deviation. *Journal of Applied Statistical Science* (2010);18(2): 121-130.
11. MacEachern S, Ozturk "O," Wolfe DA. A new ranked set sample estimator of variance. *J Royal Stat Soc B.*(2002); 64(2): 177-188.
12. McIntyre GA. A method for unbiased selective sampling using ranked sets. *Australian J Agri Res.*1952; 3(4): 385-390.
13. Mood AM, Graybill FA, Boes DC. Introduction to the Theory of Statistics (1974). McGrawM Hill Kogakusha; (1974).
14. Rousseeuw PJ, Croux C. Alternatives to the median absolute deviation. *J Am Stat Assoc.*1993; 88(424):1273-83.
15. Scheffe H. The analysis of variance. 1959. New York. (1959);331-67.
16. Shoemaker LH. Fixing the F test for equal variances. *Am Stat.* 2003;57(2):105- 14.
17. Stokes SL. Estimation of variance using judgment ordered ranked set samples. *Biometrics.* 1980; 35-42.
18. Takahasi K, Wakimoto K. On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annal Inst Stat Math.* 1968; 20(1):1-31.
19. Terpstra JT, Wang P. Confidence intervals for a population proportion based on a ranked set sample. *J Stat Comput Simul.* 2008; 78(4):351-66.