SORT 36 (1) January-June 2012, 45-68

A comparison of some confidence intervals for estimating the population coefficient of variation: a simulation study

Monika Gulhar[1], B. M. Golam Kibria[2], Ahmed N. Albatineh[3] and Nasar U. Ahmed[4]

Abstract

This paper considers several confidence intervals for estimating the population coefficient of variation based on parametric, nonparametric and modified methods. A simulation study has been conducted to compare the performance of the existing and newly proposed interval estimators. Many intervals were modified in our study by estimating the variance with the median instead of the mean and these modifications were also successful. Data were generated from normal, chi-square, and gamma distributions for CV = 0.1, 0.3, and 0.5. We reported coverage probability and interval length for each estimator. The results were applied to two public health data: child birth weight and cigarette smoking prevalence. Overall, good intervals included an interval for chi-square distributions by McKay (1932), an interval estimator for normal distributions by Miller (1991), and our proposed interval.

MSC: Primary 62F10; Secondary 62F35

Keywords: Average width, coefficient of variation, inverted coefficient of variation, confidence interval, coverage probability, simulation study, skewed distributions.

1. Introduction

The population coefficient of variation (CV) is a dimensionless (unit-free) measure of the dispersion of a probability distribution. More specifically, it is a measure of variabil-

[1] Department of Biostatistics. Florida International University. Miami. FL 33199. USA. Mgulhar@gmail.com.

[2] Department of Mathematics and Statistics. Florida International University. Miami. FL 33199. USA. kibriag@fiu.edu.

[3] Department of Biostatistics. Florida International University. Miami. FL 33199. USA. aalbatin@fiu.edu.

[4] Department of Epidemiology. Florida International University. Miami. FL 33199. USA. ahmedn@fiu.edu.
Received: April 2011
Accepted: December 2011

ity relative to the mean. This measure can be used to make comparisons across several populations that have different units of measurement. The population CV is defined as a ratio of the population standard deviation (σ) to the population mean ($\mu \neq 0$)

$$\mathrm{CV} = \frac{\sigma}{\mu} \tag{1}$$

In real life instances the population parameters σ and μ are estimated by the sample estimators s and $\bar{x}$, respectively. One obvious disadvantage arises when the mean of a variable is zero. In this case, the CV cannot be calculated. For small values of the mean, the CV will approach infinity and hence becomes sensitive to small changes in the mean. Even if the mean of a variable is not zero, but the variable contains both positive and negative values and the mean is close to zero, then the CV can be misleading.

The CV is widely used in health sciences in descriptive and inferential manners (Kelley, 2007). Some previous applications include measuring the variation in the mean synaptic response of the central nervous system (Faber and Korn, 1991) and measuring the variability in socioeconomic status and prevalence of smoking among tobacco control environments (Bernatm et al., 2009). The CV has also been used to study the impact of socioeconomic status on hospital use in New York City (Billings et al., 1993). These studies generated the CV for ambulatory care sensitive admissions for nine age cohorts and discussed the variance relative to the average number of admissions.

Perhaps, the most important use of the CV is in descriptive studies (Panichchkikosolkul, 2009). Because the CV is a unit-less measure, the variation of two or more different measurement methods can be compared to each other. The CV is used often in public health. For instance, when assessing the overall health of an individual, the CV may be useful in a comparison of variability in blood pressure measurement (mmHg) and cholesterol measurement (mg/dL). If variance were used, rather than the CV, then these two measures would not be comparable as their units of measurement differ.

Another important application of the CV is its use in the field of quality control. The inverse of the CV or ICV is equal to the signal to noise ratio (SNR), which measures how much signal has been corrupted by noise (McGibney and Smith, 1993):

$$\mathrm{SNR} = \mathrm{ICV} = \frac{\mu}{\sigma} \tag{2}$$

The CV can be used as a measure of effect or as a point estimator (Kelley, 2007). To test the significance of the CV, a hypothesis test can be conducted and a confidence interval can be generated to reject or accept the null hypothesis. Confidence intervals associated with point estimates provide more specific knowledge about the population characteristics than the p-values in the test of hypothesis (Visintainer and Tejani, 1998). A p-value only allows one to determine if results are significant or non-significant, but a confidence interval allows for the examination of an additional factor, precision. The precision of a confidence interval can be seen through the width and coverage probability of the interval. Given constant coverage, as the width of the $(1-\alpha)$ 100% confidence interval decreases, the accuracy of the estimate increases (Kelley, 2007). The coverage level is the probability that the estimated interval will capture the true CV value (Banik and Kibria, 2010).

There are various methods available for estimating the confidence interval for a population CV. For more information on the confidence interval for the CV, we refer to Koopmans et al. (1964), Miller (1991), Sharma and Krishna (1994), McKay (1932), Vangel (1996), Curto and Pinto (2009) and recently Banik and Kibria (2011), among others. The necessary sample size for estimating a population parameter is important. Therefore, determining the sample size to estimate the population CV is also important. Tables of necessary sample sizes to have a sufficiently narrow confidence intervals under different scenarios are provided by Kelly (2007).

Many researchers considered several confidence intervals for estimating the population CV. Since the studies were conducted under different simulation conditions they are not comparable as a whole. The objective of this paper is to compare several confidence interval estimators based on parametric, nonparametric and modified methods, which are developed by several researchers at several times under the same simulation conditions. Six confidence intervals that already exist in literature are considered. Additionally, we have made median and bootstrap modifications to several existing intervals in an attempt to improve the interval behaviour. Also, we have proposed our own confidence interval estimator. A simulation study is conducted to compare the performance of the interval estimators. Most of the previous simulations and comparisons found in literature were conducted under the normality assumption. Since in real life the data could be skewed we also conducted simulations under skewed distributions (chi-square and gamma). Finally, based on the simulation results, the intervals with high coverage probability and small width were recommended for practitioners.

The organization of this paper is as follows: In Section 2 we present the proposed confidence intervals. In Section 3, the simulation technique and results are provided and discussed. Two real life data are analyzed in Section 4. Finally, some concluding remarks are presented in Section 5. Due to space limitations, only graphs representing simulations for the normal distribution are presented but the full simulation results are presented in the Appendix.

2. Statistical methodology

Let $x_1, x_2, x_3, \ldots, x_n$ be an independently and identically distributed (iid) random sample of size n from a distribution with finite mean, μ, and finite variance, σ^2. Let $\bar{x}$ be the sample mean and s be the sample standard deviation. Then $\widehat{\mathrm{CV}} = s/\bar{x}$ would be the estimated value of the population CV. We want to find the $(1-\alpha)100\%$ confidence intervals for the population CV. In this section we will review some existing interval estimators and propose some new interval estimators for CV. A total of 15 intervals will be considered.

2.1. The existing confidence interval estimators

Six existing (parametric and nonparametric) confidence intervals for the CV are reviewed in this section.

2.1.1. Parametric confidence intervals

1. Miller's (1991) confidence interval for normal distribution (referred to as Mill). Miller showed that $\frac{s}{\bar{x}}$ approximates an asymptotic normal distribution with mean $\frac{\sigma}{\mu}$ and variance $m^{-1}(\frac{\sigma}{\mu})^2\left(0.5+(\frac{\sigma}{\mu})^2\right)$. Then, the $(1-\alpha)100\%$ approximate confidence interval for the population $\mathrm{CV}=\frac{\sigma}{\mu}$ is given by

$$\frac{s}{\bar{x}} \pm Z_{\frac{\alpha}{2}}\sqrt{m^{-1}\left(\frac{s}{\bar{x}}\right)^2\left(0.5+\left(\frac{s}{\bar{x}}\right)^2\right)} \tag{3}$$

which can be expressed as

$$\frac{s}{\bar{x}} - Z_{\frac{\alpha}{2}}\sqrt{m^{-1}\left(\frac{s}{\bar{x}}\right)^2\left(0.5+\left(\frac{s}{\bar{x}}\right)^2\right)} < \mathrm{CV} < \frac{s}{\bar{x}} + Z_{\frac{\alpha}{2}}\sqrt{m^{-1}\left(\frac{s}{\bar{x}}\right)^2\left(0.5+\left(\frac{s}{\bar{x}}\right)^2\right)} \tag{4}$$

where $m=n-1, \bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ is the sample mean, $s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ is the sample standard deviation and Z_α is the upper $(1-\alpha)100^{th}$ percentile of the standard normal distribution.

2. Sharma and Krishna's (1994) confidence interval for inverted CV (S&K): This confidence interval is given by

$$\frac{\bar{x}}{s}+\frac{\phi^{-1}_{(\alpha/2)}}{\sqrt{n}} < \frac{1}{\mathrm{CV}} < \frac{\bar{x}}{s}-\frac{\phi^{-1}_{(\alpha/2)}}{\sqrt{n}} \tag{5}$$

where ϕ cumulative standard normal distribution. Therefore, the $(1-\alpha)100\%$ confidence interval for the population CV is given by

$$\left(\frac{\bar{x}}{s}-\frac{\phi^{-1}_{(\alpha/2)}}{\sqrt{n}}\right)^{-1} < \mathrm{CV} < \left(\frac{\bar{x}}{s}+\frac{\phi^{-1}_{(\alpha/2)}}{\sqrt{n}}\right)^{-1} \tag{6}$$

3. Curto and Pinto's (2009) iid assumption (C&P):

$$\frac{s}{\bar{x}} - Z_{(\alpha/2)}\sqrt{\frac{1}{n}\left(\left(\frac{s}{\bar{x}}\right)^4 + \frac{1}{2}\left(\frac{s}{\bar{x}}\right)^2\right)} < \text{CV} < \frac{s}{\bar{x}} + Z_{(\alpha/2)}\sqrt{\frac{1}{n}\left(\left(\frac{s}{\bar{x}}\right)^4 + \frac{1}{2}\left(\frac{s}{\bar{x}}\right)^2\right)} \tag{7}$$

2.1.2. Confidence intervals based on chi-square distribution

1. McKay's (1932) confidence interval for chi-square distribution (McK):

$$\left(\frac{s}{\bar{x}}\right)\sqrt{\left(\frac{\chi^2_{\nu,1-\alpha/2}}{\nu+1} - 1\right)\left(\frac{s}{\bar{x}}\right)^2 + \frac{\chi^2_{\nu,1-\alpha/2}}{\nu}} < \text{CV} < \left(\frac{s}{\bar{x}}\right)\sqrt{\left(\frac{\chi^2_{\nu,\alpha/2}}{\nu+1} - 1\right)\left(\frac{s}{\bar{x}}\right)^2 + \frac{\chi^2_{\nu,\alpha/2}}{\nu}} \tag{8}$$

   where $\nu = n-1$ and $\chi^2_{\nu,\alpha}$ is the 100α-th percentile of a chi-square distribution with ν degrees of freedom.

2. Modified McKay (1996) confidence interval (MMcK): Vangel (1996) modified McKay's original (1932) interval:

$$\left(\frac{s}{\bar{x}}\right)\sqrt{\left(\frac{\chi^2_{(\nu,1-\alpha/2)}+2}{\nu+1} - 1\right)\left(\frac{s}{\bar{x}}\right)^2 + \frac{\chi^2_{\nu,1-\alpha/2}}{\nu}} < \text{CV} < \left(\frac{s}{\bar{x}}\right)\sqrt{\left(\frac{\chi^2_{(\nu,\alpha/2)}+2}{\nu+1} - 1\right)\left(\frac{s}{\bar{x}}\right)^2 + \frac{\chi^2_{\nu,\alpha/2}}{\nu}} \tag{9}$$

3. Panichkitkosolkul's (2009) confidence interval (Panich): Panichkitkosokul further modified the Modified McKay interval by replacing the sample CV with the maximum likelihood estimator for a normal distribution, $\tilde{k}$

$$\tilde{k}\sqrt{\left(\frac{\chi^2_{(\nu,1-\alpha/2)}+2}{\nu+1} - 1\right)\left(\tilde{k}\right)^2 + \frac{\chi^2_{\nu,1-\alpha/2}}{\nu}} < \text{CV} < \tilde{k}\sqrt{\left(\frac{\chi^2_{(\nu,\alpha/2)}+2}{\nu+1} - 1\right)\left(\tilde{k}\right)^2 + \frac{\chi^2_{\nu,\alpha/2}}{\nu}} \tag{10}$$

   where $\tilde{k} = \frac{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\sqrt{n}\bar{x}}$

2.2. Median modifications of existing intervals

Kibria (2006) and Shi and Kibria (2007) claimed that for a skewed distribution, the median describes the centre of the distribution better than the mean. Thus, for skewed data it makes more sense to measure sample variability in terms of the median rather than the mean. Following Shi and Kibria (2007), the $(1-\alpha)100\%$ CI for the CV are obtained for four of the existing estimators and provided below. These median modifications are made in attempt to improve the performance of the original intervals. The intervals selected for modification represent parametric and non-parametric estimators.

1. Median Modified Miller Estimator (Med Mill):

$$\frac{\tilde{s}}{\tilde{x}} - Z_{\alpha/2}\sqrt{m^{-1}\left(\frac{\tilde{s}}{\tilde{x}}\right)^2\left(0.5+\left(\frac{\tilde{s}}{\tilde{x}}\right)^2\right)} < \mathrm{CV} < \frac{\tilde{s}}{\tilde{x}} + Z_{\alpha/2}\sqrt{m^{-1}\left(\frac{\tilde{s}}{\tilde{x}}\right)^2\left(0.5+\left(\frac{\tilde{s}}{\tilde{x}}\right)^2\right)} \tag{11}$$

where $\tilde{s} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\tilde{x})^2}$ and $\tilde{x}$ is the sample median.

2. Median Modification of McKay (Med McK):

$$\left(\frac{\tilde{s}}{\tilde{x}}\right)\sqrt{\left(\frac{\chi^2_{\nu,1-\alpha/2}}{\nu+1}-1\right)\left(\frac{\tilde{s}}{\tilde{x}}\right)^2+\frac{\chi^2_{\nu,1-\alpha/2}}{\nu}} < \mathrm{CV} < \left(\frac{\tilde{s}}{\tilde{x}}\right)\sqrt{\left(\frac{\chi^2_{\nu,\alpha/2}}{\nu+1}-1\right)\left(\frac{\tilde{s}}{\tilde{x}}\right)^2+\frac{\chi^2_{\nu,\alpha/2}}{\nu}} \tag{12}$$

3. Median Modification of Modified McKay (Med MMcK):

$$\left(\frac{\tilde{s}}{\tilde{x}}\right)\sqrt{\left(\frac{\chi^2_{\nu,1-\alpha/2}+2}{\nu+1}-1\right)\left(\frac{\tilde{s}}{\tilde{x}}\right)^2+\frac{\chi^2_{\nu,1-\alpha/2}}{\nu}} < \mathrm{CV} < \left(\frac{\tilde{s}}{\tilde{x}}\right)\sqrt{\left(\frac{\chi^2_{\nu,\alpha/2}+2}{\nu+1}-1\right)\left(\frac{\tilde{s}}{\tilde{x}}\right)^2+\frac{\chi^2_{\nu,\alpha/2}}{\nu}} \tag{13}$$

4. Median Modified Curto and Pinto(2009) iid assumption (Med C&P):

$$\frac{\tilde{s}}{\tilde{x}} - Z_{(\alpha/2)}\sqrt{\frac{\left(\frac{\tilde{s}}{\tilde{x}}\right)^4+0.5\left(\frac{\tilde{s}}{\tilde{x}}\right)^2}{n}} < \mathrm{CV} < \frac{\tilde{s}}{\tilde{x}} + Z_{(\alpha/2)}\sqrt{\frac{\left(\frac{\tilde{s}}{\tilde{x}}\right)^4+0.5\left(\frac{\tilde{s}}{\tilde{x}}\right)^2}{n}} \tag{14}$$

2.3. Bootstrap confidence intervals

Bootstrapping is a commonly used computer-intensive, nonparametric tool (introduced by Efron, 1979), which is used to make inference about a population parameter when there are no assumptions regarding the underlying population available. It will be especially useful because unlike other methods, this technique does not require any assumptions to be made about the underlying population of interest (Banik and Kibria, 2009). Therefore, bootstrapping can be applied to all situations. This method is implemented by simulating an original data set then randomly selecting variables several times with replacement to estimate the CV. The accuracy of the bootstrap CI depends on the number of bootstrap samples. If the number of bootstrap samples is large enough, the CI may be very accurate. We will consider four different bootstrap methods which are summarized in this section.

Let $X^{(*)} = X_1^{(*)}, X_2^{(*)}, X_3^{(*)}, \ldots, X_n^{(*)}$ where the i^{th} sample is denoted $X^{(i)}$ for $i = 1, 2, \ldots, B$ and B is the number of bootstrap samples. Efron (1979) suggests using a minimum value of $B = 1000$. Following Banik and Kibria (2009), we will consider four different bootstrap methods for estimating the population CV which are summarized in this section.

2.3.1. Non-parametric bootstrap CI (NP BS)

The CV for all bootstrap samples is computed and then ordered as follows:

$$\mathrm{CV}^*_{(1)} \leq \mathrm{CV}^*_{(2)} \leq \mathrm{CV}^*_{(1)} \cdots \mathrm{CV}^*_{(B)}$$

The lower confidence level (LCL) and upper confidence level (UCL) for the population CV is then

$$\mathrm{LCL} = \mathrm{CV}^*_{[(\alpha/2)B]} \quad \text{and} \quad \mathrm{UCL} = \mathrm{CV}^*_{[(1-\alpha/2)B]}$$

Therefore, in a case where $B = 1000$, LCL = the 25^{th} bootstrap sample and UCL = the 975^{th} bootstrap sample.

2.3.2. Bootstrap t-approach (BS)

Following Banik and Kibria (2009), we will propose a version called the bootstrap-t, defined as

$$\mathrm{LCL} = \widehat{\mathrm{CV}} + T^*_{(n-1),\alpha/2} S_{\widehat{\mathrm{CV}}} \quad \text{and} \quad \mathrm{UCL} = \widehat{\mathrm{CV}} + T^*_{(n-1),1-\alpha/2} S_{\widehat{\mathrm{CV}}}$$

where $\widehat{\mathrm{CV}}$ is the sample CV, $T^*_{\alpha/2}, T^*_{1-\alpha/2}$ are the $(\alpha/2)^{th}$ and $(1-\alpha/2)^{th}$ sample quantiles of

$$T_i^* = \frac{\mathrm{CV}_i^* - \overline{\overline{\mathrm{CV}}}}{\widehat{\sigma}_{\mathrm{CV}}}$$

where

$$\widehat{\sigma}_{\mathrm{CV}} = \sqrt{\frac{1}{B}\sum_{i=1}^{B}\left(\mathrm{CV}_i^* - \overline{\overline{\mathrm{CV}}}\right)^2} \quad \text{and} \quad \overline{\overline{\mathrm{CV}}} = \frac{1}{B}\sum_{i=1}^{B}\mathrm{CV}_i^*$$

is a bootstrap CV.

2.3.3. Modified median Miller based on critical value from bootstrap samples (BSMill)

$$\frac{\tilde{s}}{\bar{x}}+T^*_{\alpha/2}\sqrt{m^{-1}\left(\frac{\tilde{s}}{\bar{x}}\right)^2\left(0.5+\left(\frac{\tilde{s}}{\bar{x}}\right)^2\right)} < \mathrm{CV} < \frac{\tilde{s}}{\bar{x}}+T^*_{1-(\alpha/2)}\sqrt{m^{-1}\left(\frac{\tilde{s}}{\bar{x}}\right)^2\left(0.5+\left(\frac{\tilde{s}}{\bar{x}}\right)^2\right)} \tag{15}$$

2.3.4. Modified median Curto and Pinto (2009) based on BS sample (BS C&P)

$$\frac{\tilde{s}}{\bar{x}}-T^*_{(\alpha/2)}\sqrt{\frac{\left(\frac{\tilde{s}}{\bar{x}}\right)^4+0.5\left(\frac{\tilde{s}}{\bar{x}}\right)^2}{n}} < \mathrm{CV} < \frac{\tilde{s}}{\bar{x}}+T^*_{1-(\alpha/2)}\sqrt{\frac{\left(\frac{\tilde{s}}{\bar{x}}\right)^4+0.5\left(\frac{\tilde{s}}{\bar{x}}\right)^2}{n}} \tag{16}$$

where $T^*_{\alpha/2}$ and $T^*_{1-\alpha/2}$ are the $(\alpha/2)^{th}$ and $(1-\alpha/2)^{th}$ sample quantiles of $T^*_i = \frac{(\mathrm{CV}^*_i-\overline{\overline{\mathrm{CV}}})}{\hat{\sigma}_{\mathrm{CV}}}$

2.4. Proposed confidence interval based on estimator of σ^2

Another parametric confidence interval that was compared in the simulation was one based on the known formula for calculating the confidence interval for σ^2. Let $x_1, x_2, x_3, \ldots, x_n$ be an iid random sample which is normally distributed with finite mean μ and variance σ^2. Then the $(1-\alpha)100\%$ CI for σ^2 is

$$\frac{(n-1)s^2}{\chi^2_{\nu,1-\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\nu,\alpha/2}} \tag{17}$$

Assuming that $\mu \neq 0$, dividing this interval by μ^2 results in

$$\frac{(n-1)s^2}{\left(\chi^2_{\nu,1-\alpha/2}\right)\mu^2} < \left(\frac{\sigma}{\mu}\right)^2 < \frac{(n-1)s^2}{\left(\chi^2_{\nu,\alpha/2}\right)\mu^2} \tag{18}$$

Since μ is not known, we can replace it by the unbiased estimator of μ which is $\bar{x}$ resulting in

$$\frac{(n-1)s^2}{\left(\chi^2_{\nu,1-\alpha/2}\right)\bar{x}^2} < \mathrm{CV}^2 < \frac{(n-1)s^2}{\left(\chi^2_{\nu,\alpha/2}\right)\bar{x}^2} \tag{19}$$

Taking the square root results in the final proposed interval estimator given by

$$\frac{\sqrt{n-1}\,\widehat{\mathrm{CV}}}{\sqrt{\chi^2_{\nu,1-\alpha/2}}} < \mathrm{CV} < \frac{\sqrt{n-1}\,\widehat{\mathrm{CV}}}{\sqrt{\chi^2_{\nu,\alpha/2}}} \tag{20}$$

3. Simulation study

Many researchers have considered several confidence intervals for estimating the population CV. However, these studies were all conducted under different simulation conditions and therefore they are not comparable as a whole. In this study we considered 15 useful confidence intervals (six existing intervals, eight modified intervals, and one proposed interval) for estimating the population CV and compared them under the same simulation conditions. A Monte-Carlo simulation was conducted using the R statistical software (2010) version 2.10.1 to compare the performance of the interval estimators. The performance of the estimators was considered for various CV values, sample sizes, and distributions.

Table 1: *The CV and skewness of data from normal, chi-square, and gamma distributions.*

Distribution	CV	Skewness
$N(\mu, \sigma)$	$\frac{\sigma}{\mu}$	0
chi-square (ν)	$\sqrt{\frac{2}{\nu}}$	$2\sqrt{\frac{2}{\nu}}$
gamma (α,2)	$\frac{1}{\sqrt{\alpha}}$	$\frac{2}{\sqrt{\alpha}}$

Table 2: *Estimated lower and upper confidence limits and corresponding widths for all proposed methods for Example 1 ($n = 189$).*

Method	Lower Confidence Limit	Upper Confidence Limit	Width
Mill	0.223719	0.277478	0.053759
S&K	0.223659	0.277365	0.053706
McK	0.223073	0.276620	0.053546
MMcK	0.221061	0.274088	0.053027
C&P	0.239134	0.256633	0.017499
Panich	0.230751	0.286434	0.055683
Proposed	0.224879	0.275406	0.050527
Med Mill	0.223940	0.277752	0.053812
MedMcK	0.223878	0.277643	0.053765
MedMMcK	0.221277	0.274362	0.053086
Med C&P	0.221347	0.274292	0.052945
NP BS	0.220954	0.274950	0.053996
P BS	0.219723	0.274653	0.054930
BS Mill	0.220414	0.273982	0.053568
BS C&P	0.220486	0.273912	0.053426

3.1. Simulation technique

To study the behaviour of small and large sample sizes, we used $n = 15, 25, 50, 100$. To examine the behaviour of the intervals when the sample size is increased even further, a second simulation was conducted using $n = 500$. A random sample was generated with specific parameters from a normal distribution, chi-square distribution and gamma distribution. The probability density functions of the distributions are given below.

- Normal distribution, with $\mu = 10$ and $\sigma = 1, 3, 5$ where

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-10}{\sigma}\right)^2}, \quad -\infty < x < \infty, \quad \sigma > 0$$

- Chi-square distribution, with degrees of freedom (df) $v = 200, 22, 8$, where

$$f(x) = \frac{1}{2^{\frac{v}{2}}\Gamma(v/2)} e^{-\frac{x}{2}} x^{\frac{v}{2}-1}, \quad x \geq 0$$

- Gamma distribution with $\alpha = 100, 11.11, 4$ and $\beta = 2$, where

$$f(x) = \frac{1}{\sqrt{\alpha}} \frac{e^{-\frac{x}{2}} x^{\alpha-1}}{2^{\alpha}}, \quad x > 0, \alpha > 0$$

The CV was calculated for each distribution type by utilizing the equations in Table 1. Thus, based on the CV formula presented in Table 3.1, all three distribution types consider CV $= 0.10, 0.30, 0.50$. The number of simulation replications (M) was 2000 for each case. For intervals that utilize the bootstrapping technique, 1000 bootstrap samples (B) are used for each n.

Coverage probability and width of the interval were measured for each case. The most common 95% confidence interval ($\alpha = 0.05$) is used for measuring the confidence level. The coverage probability is calculated by counting the number of times the true CV is captured between the upper and lower limits. Generally, since the sampling distribution of the sample CV is approximately normally distributed, see Miller (1991), as the sample size (number of sample CV's) increases the distribution becomes more symmetric and it is expected for the coverage probability to approach $(1-\alpha)$. When $\alpha = 0.05$, an interval that has perfect performance in terms of coverage probability will capture the true CV between the upper and lower limits 95% of the time. The coverage probability is an excellent method for evaluating the success of a particular interval in capturing the true parameter. An interval width is calculated by subtracting a lower limit from an upper limit. A smaller width is better because it means that the true CV is captured within a smaller span and the results are more precise. In cases where the coverage probability is comparable, the interval length is especially important because the smaller width will give a more accurate and precise result. The simulated coverage probabilities and interval widths for the normal, chi-square and

Table 3: *Estimated lower and upper confidence limits and corresponding widths for all proposed methods for Example 2 ($n = 24$).*

Method	Lower Confidence Limit	Upper Confidence Limit	Width
Mill	0.181510	0.340507	0.158997
S&K	0.181261	0.338874	0.157613
McK	0.177526	0.331398	0.153872
MMcK	0.163988	0.307669	0.143680
C&P	0.215497	0.260397	0.044900
Panich	0.186213	0.350015	0.163802
Proposed	0.165501	0.306156	0.140655
Med Mill	0.183289	0.330811	0.147522
MedMcK	0.187048	0.350897	0.163849
MedMMcK	0.186665	0.349749	0.163084
Med C&P	0.168763	0.317286	0.148523
NP BS	0.170327	0.315722	0.145396
P BS	0.188882	0.340905	0.152023
BS Mill	0.121866	0.305541	0.183675
BS C&P	0.136045	0.319795	0.183750

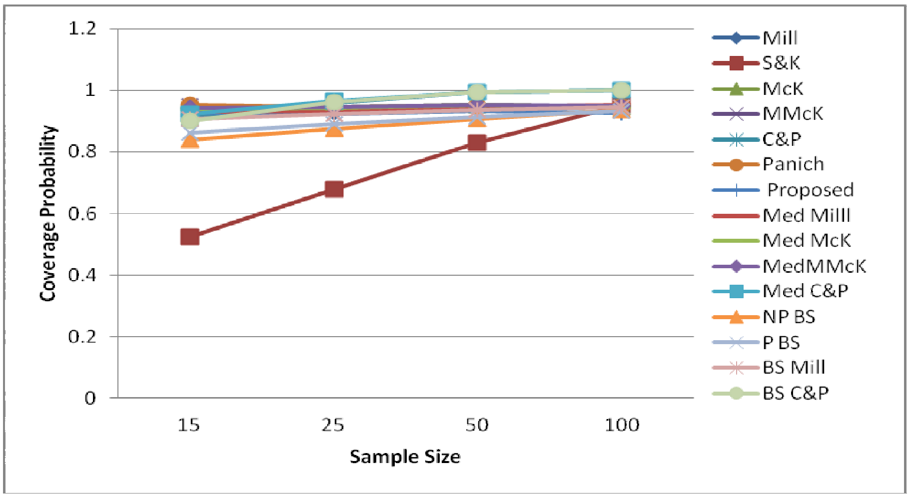

Figure 1: *Sample size vs. coverage probability for normal distribution when* CV = *0.30.*

gamma distributions are presented in Tables 4, 5 and 6 of the appendix, respectively, and for $n = 500$ of all distributions are presented in Table 7 of the appendix. Each table gives results for the various sample sizes and CV values previously mentioned. Due to space limitation simulation results of the normal distribution are graphically displayed in Figure 1 through Figure 4, but full simulation results are given in the Appendix.

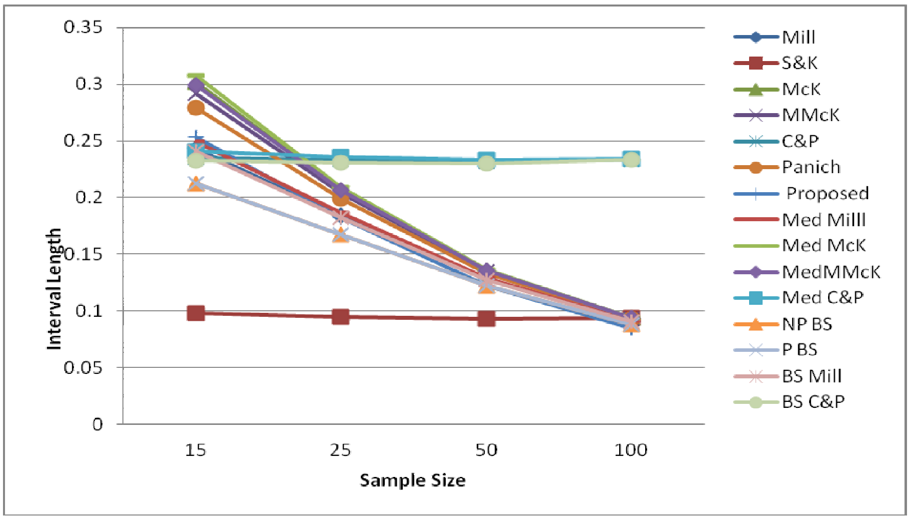

Figure 2: *Sample size vs. interval length for normal distribution when CV = 0.30.*

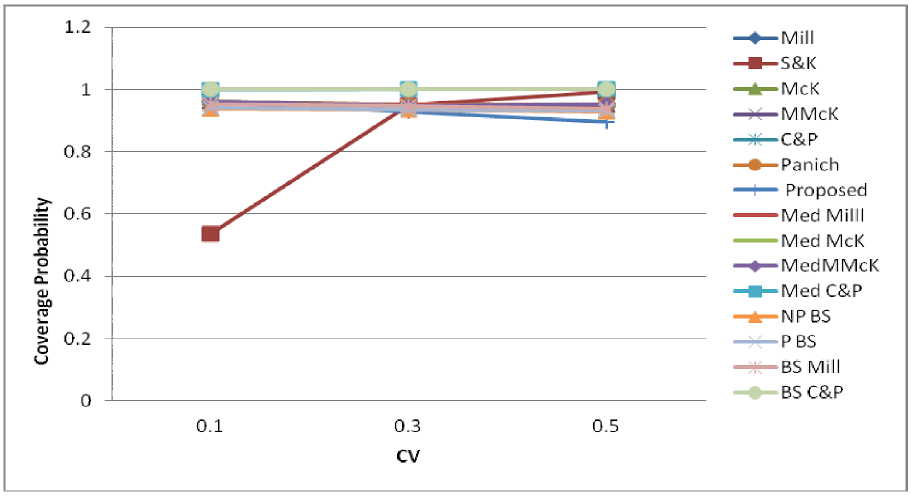

Figure 3: *Coefficient of variation vs. coverage probability for normal distribution when* $n = 100$.

3.2. Simulation results

A main trend that was noted throughout all distributions was the net increase in coverage probability from $n = 15$ to $n = 100$. The results will be discussed below by distribution type. To aid in visualization of results, multiple graphs have been referred to.

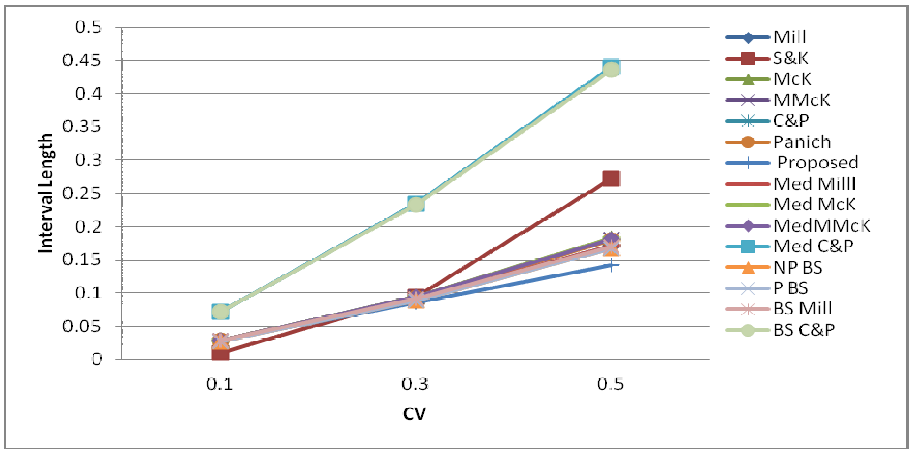

Figure 4: *Coefficient of variation vs. interval length for normal distribution when* $n = 100$.

3.2.1. Normal distribution

Table 4 and Figures 1–4 present simulated results for the normal distribution. Figure 1 displays sample size vs. coverage probability for all tested intervals when CV = 0.30. This figure clearly shows that S&K had the weakest overall performance with its lowest coverage probability at 21% (Table 4). However, S&K's performance improves as sample size increases. By $n = 100$, almost all intervals are performing at a similar level (Figure 1). All C&P intervals (C&P, Med C&P, and BS C&P) over exceeded the expected coverage probability of 95% and reached 100% and are clear outliers. Figure 2 displays the sample size vs. interval length for all intervals when CV = 0.30. All C&P intervals seem to remain stable and large throughout all sample sizes. The S&K interval also remains stable; however, the values are, as desired, low. All other intervals follow a similar trend: as sample size increases, interval length decreases. When $n = 100$, all intervals are performing similarly, however the C&P intervals are clear outliers. Figure 3 represents the CV vs. coverage probability for normal distribution when $n = 100$. All intervals seem to fall between a coverage range of .9 and 1 however, when CV = 0.10 the S&K interval performs very poorly.

The largest sample size presented in Tables 4, 5, and 6 is $n = 100$. For instances where one wants to consider an even larger sample size, Table 7 presents the behaviour of each interval for all distribution types when $n = 500$. Overall, all intervals seem to perform at a similar level when $n = 500$. One key difference that should be noted is that when sample size in increased to 500, there is a drastic improvement in the performance of S&K.

3.2.2. Skewed distributions

The simulated results for chi-square distribution are presented in Table 5 which indicates that many of the same trends as seen in the normal distribution. Similar to the normal distribution, the S&K interval had the weakest overall performance for the chi-square distribution. Table 5 shows that as sample size increases, coverage probability for all intervals has a general increasing trend. The S&K interval is notably lacking in performance relatively to other intervals. Table 5 shows that as sample size increases, the interval length decreases for most of the intervals. Similar to the results seen for the normal distribution, all C&P intervals seem to remain stable and large throughout all sample sizes and do not follow the decreasing trend. The S&K interval also did not follow the trend as it remained low throughout all sample sizes. Table 5 indicates that with the exception of S&K, all intervals had coverage probabilities between 0.9 and 1 for all CV values. S&K performed much lower than the other intervals when CV $= 0.10$. The same trend was seen for the normal distribution. Table 5 shows that as the CV value increases, all interval lengths become increasingly wider. The widths are narrowest when CV $= 0.10$ and widest when CV $= 0.50$. The C&P intervals have the greatest length for all CV values. When CV $= 0.50$, S&K also has a notably higher interval length than the other intervals and the narrowest width is observed in the BS and PBS intervals.

The results for gamma distribution are presented in Table 6. Although the gamma distribution has a greater skewness than the chi-square distribution, this difference did not have a major effect on the trends seen between the two distribution types. All trends in the gamma distribution are comparable to results from the chi-square distribution because the two distribution types are related. When sample size is increased to 500 for skewed distributions, all intervals perform at a similar level (Table 7). The weakest interval when $n = 500$ is C&P. C&P has its lowest coverage probability of 21% when CV $= 0.10$ in the gamma distribution.

4. Application

To illustrate the findings of the paper, some real life health data are analyzed in this section.

4.1. Example 1: Birth weight data

The first data set was obtained from Hosmer and Lemeshow (2000), which was collected from the Baystate Medical Center, in Springfield, Massachusetts (University of Massachusetts Amherst, 2000). Child birth weight data were collected from 189 women. A baby weighing less than 2500 grams was defined as a "low birth weight" child. Among them, 59 women had low birth weight babies and 130 women had normal birth weight babies. The average birth weight was 2944.66 grams, with a standard deviation of 729.022 grams. The coefficient of variation for the low birth data is 0.248. The histogram of the data is presented in Figure 5, which showed a mound shaped distribution. The Shapiro-Wilk test (W = 0.9925, p-value = 0.4383) supported that the data follow a normal distribution.

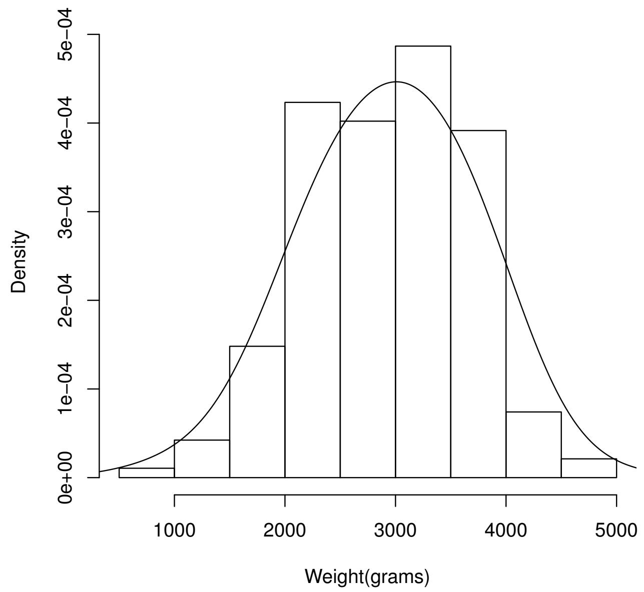

Figure 5: *Histogram of birth weight of babies from 189 women, with modelled normal distribution.*

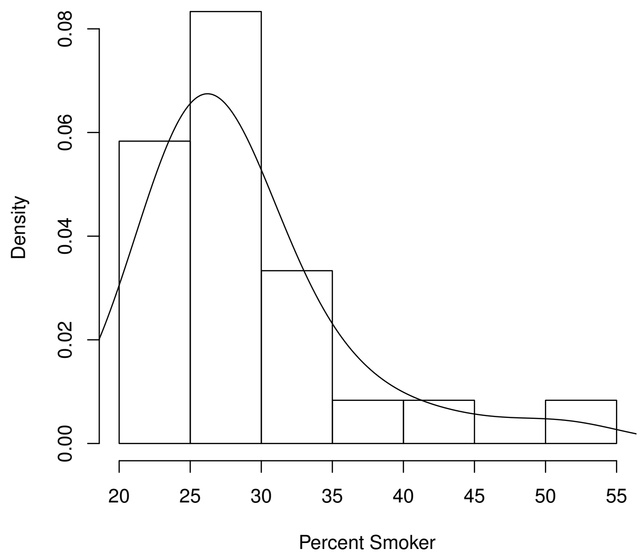

Figure 6: *Histogram of the percent of smokers in the United States for 24 selected years, with modelled gamma distribution.*

Using SPSS (2008), the Kolmogorov-Smirnov (ks) goodness of fit test result (ks = 0.043, p-value = 0.200[1]) for a hypothesized normal distribution with mean $\mu = 2945$ and standard deviation $\sigma = 755$, indicates that the data follow a normal distribution with mean of 2945 gm and standard deviation of 755 gm. The corresponding population coefficient of variation, CV $= 755/2945 = 0.2564$. The resulting 95% confidence intervals and corresponding widths for all proposed intervals are reported in Table 2. Results showed that all confidence intervals covered the hypothesize population CV 0.2564. We note that the C&P interval has the shortest width followed by our proposed interval. Panich has the largest width. Based on the overall simulation results presented and Figure 3, we may suggest S&K or the proposed one to be the interval of choice for this example. In fact the proposed interval performed better than the majority of the other interval estimators.

4.2. Example 2: Cigarette smoking among men, women, and high school students: United States, 1965-2007

The data set for this example were obtained from the Center for Disease Control and Prevention (CDC, 2009). The CDC used the National Health Interview Survey and the Youth Risk Behavior Survey to compile the data. Adults were classified as cigarette smokers if they smoked 100 cigarettes in their lifetime and if now smoke daily or occasionally.

A high school student was categorized as a smoker if he or she had even one cigarette in the past month. The data is a compilation of cigarette smoking among adult and high school population for selected years between 1965 and 2007. The total sample size for the data is $n = 24$ selected years between 1965 and 2007. The mean percent of smokers is 28.9%, with a standard deviation of 6.8%. The coefficient of variation for the data is 0.236. The histogram of the data is presented in Figure 6, which shows a right skewed distribution. Using SAS (2008), the Kolmogorov-Smirnov (ks) goodness of fit test result (ks = 0.188, p-value > 0.25) for a hypothesized gamma distribution with shape parameter, $\alpha = 17.6$ and scale parameter $\beta = 1.6$, indicates that the data follow a gamma distribution with $\alpha = 17.6$ and $\beta = 1.6$. For more information on goodness of fit for the gamma distribution, see D'Agostino and Stephens (1986). For this example, the corresponding population coefficient of variation, CV $= \frac{1}{\sqrt{17.6}} = 0.238$. The resulting 95% confidence intervals and corresponding widths for all proposed intervals are reported in Table 3. Results showed that all confidence intervals covered the hypothesize population CV of 0.238. Based on the simulation results presented, we would have expected the S&K interval to have the narrowest width,

1. Lillifors significance correction is applied. This is a lower bound of the true significance.

however, in this example the narrowest width was observed in the original C&P interval. S&K's interval had results that were very comparable to all other intervals, but it was not the best. This example indicates that the C&P bootstrap interval is the narrowest, a finding that is confirmed by our simulation results.

5. Concluding remarks

We have considered several confidence intervals for estimating the population CV. A simulation study has been conducted to compare the performance of the estimators. For simulation purposes, we have generated data from both symmetric (normal) and skewed distributions (chi-square and gamma) to see the performance of the interval estimators. After thorough examination of all individual intervals, and each overall simulation condition, it can be concluded that the intervals that performed well are: Mill, McK, MMcK, Panich, MedMill, MedMcK, MedMMcK, and our proposed interval. The S&K interval was not a good estimator for small sample sizes and had the weakest performance in terms of coverage probability. However, its performance is comparable to other intervals when the sample size is large. All C&P intervals did not perform well for large sample sizes as the interval length for these estimators was too high relative to other estimators. This wide interval length is not desired as it indicates that the results are imprecise. It is known that for symmetric distributions the mean equals the median, but as pointed out by a reviewer, the median modified intervals of the CV for skewed distributions works well for moderate sample sizes, but for a very large sample size, the point estimate of the CV using the median will be very close but doesn't converge to the true population CV. To see the large sample behaviour, we have generated data from normal, chi-square and gamma distribution for $n = 500$ and presented the coverage probability and average widths in Table 7. Clearly, Table 7 shows that the coverage probabilities for all estimators are close to the nominal size of 0.95. For large sample size any of the intervals except S&K could be used. We do not encourage the researchers to use bootstrap confidence intervals as their performances are not significantly better than others and they are very time consuming. However, if some researchers are very conservative about the assumptions of the distributions and willing to accept the extensive computation they might consider the bootstrap methods. Most importantly, in many instances, our newly proposed interval produced the best interval length, especially as CV increased from 0.1 to 0.5. Higher values of CV indicate more variability in the data, a characteristic that is often seen in health sciences where sample sizes are frequently small. Because our proposed interval performed well for the higher values of CV, it will be a good interval to use for many health science data. We also recommend using this interval over other intervals because it is a very user-friendly interval in calculation.

Acknowledgment

The authors would like to thank the editor and two anonymous reviewers for their valuable comments and suggestions that greatly improved the content and presentation of an earlier draft of the paper.

Appendix

The appendix contains Table 4 through Table 7.

Table 4: *Estimated coverage probabilities and widths of the intervals using normal distribution ($M = 2000$).*

	Mill^2	S&K	McK	MMcK	C&P	Panich	Prop	Med Mill	Med McK	Med MMcK	MedC&P	NPBS	PBS	BSMill	BS C&P
N=15, CV=.1															
Coverage	0.9235	0.212	0.9535	0.9535	0.9155	0.951	0.952	0.9325	0.9445	0.9445	0.9265	0.8395	0.87	0.909	0.9025
Length	0.0746	0.0104	0.0857	0.0854	0.072066	0.0824	0.0842	0.0759	0.0871	0.0869	0.073321	0.0645	0.0645	0.0736	0.071132
N=15, CV=.3															
Coverage	0.9165	0.5235	0.9465	0.9505	0.912	0.9525	0.928	0.9295	0.939	0.9435	0.9245	0.84	0.863	0.9075	0.9015
Length	0.2437	0.0978	0.3015	0.2917	0.235397	0.2791	0.2534	0.2491	0.3072	0.2989	0.240624	0.2122	0.2125	0.2411	0.232897
N=15, CV=.5															
Coverage	0.918	0.7265	0.952	0.9545	0.9115	0.9535	0.8915	0.9295	0.949	0.954	0.922	0.865	0.886	0.9115	0.906
Length	0.4678	0.2948	0.8265	0.689	0.451932	0.6405	0.4244	0.4788	0.8401	0.7156	0.462575	0.424	0.4239	0.4631	0.447386
N=25, CV=.1															
Coverage	0.9305	0.2785	0.9495	0.95	0.964	0.952	0.949	0.9345	0.945	0.945	0.9675	0.877	0.895	0.925	0.9565
Length	0.0565	0.0101	0.0611	0.061	0.071406	0.0598	0.0603	0.0571	0.0618	0.0617	0.072236	0.0513	0.0512	0.0559	0.070708
N=25, CV=.3															
Coverage	0.929	0.6775	0.945	0.944	0.9605	0.945	0.9235	0.934	0.9435	0.9445	0.9645	0.8745	0.89	0.922	0.962
Length	0.1843	0.0948	0.207	0.2039	0.2331	0.1989	0.1822	0.1867	0.2094	0.2068	0.236185	0.1676	0.1675	0.1827	0.231161
N=25, CV=.5															
Coverage	0.918	0.824	0.946	0.9515	0.953	0.95	0.8795	0.93	0.936	0.943	0.9585	0.8825	0.892	0.9145	0.956
Length	0.3503	0.2812	0.4454	0.422	0.443089	0.407	0.3043	0.3557	0.4505	0.4304	0.449977	0.3234	0.3232	0.3476	0.439721
N=50, CV=.1															
Coverage	0.9365	0.374	0.9475	0.9475	0.995	0.9465	0.9455	0.9395	0.9475	0.9475	0.995	0.908	0.911	0.9365	0.9945
Length	0.0398	0.0102	0.0414	0.0413	0.071923	0.0409	0.0409	0.04	0.0416	0.0416	0.072323	0.0377	0.0377	0.0395	0.071469
N=50, CV=.3															
Coverage	0.94	0.829	0.9525	0.952	0.994	0.9495	0.932	0.9445	0.952	0.9515	0.9955	0.909	0.9145	0.9365	0.993
Length	0.1281	0.0929	0.1355	0.1347	0.231613	0.1331	0.1222	0.1289	0.1362	0.1355	0.233008	0.122	0.122	0.1273	0.230112
N=50, CV=.5															
Coverage	0.934	0.9535	0.9455	0.945	0.9925	0.9495	0.8855	0.9385	0.943	0.945	0.9935	0.908	0.921	0.938	0.992
Length	0.2459	0.2794	0.2755	0.2703	0.444506	0.2662	0.2065	0.2478	0.2771	0.2726	0.447815	0.2365	0.2368	0.245	0.442756
N=100, CV=.1															
Coverage	0.954	0.5355	0.961	0.961	0.9995	0.9615	0.96	0.956	0.963	0.963	0.9995	0.938	0.9405	0.953	1
Length	0.0281	0.0102	0.0286	0.0286	0.07217	0.0285	0.0283	0.0282	0.0287	0.0287	0.072372	0.0271	0.0271	0.028	0.071852
N=100, CV=.3															
Coverage	0.9495	0.951	0.9475	0.9475	1	0.9485	0.928	0.952	0.947	0.948	1	0.935	0.9345	0.946	1
Length	0.091	0.094	0.0937	0.0934	0.233686	0.0928	0.0851	0.0912	0.0939	0.0937	0.23442	0.0881	0.0883	0.0907	0.23298
N=100, CV=.5															
Coverage	0.943	0.992	0.951	0.9505	1	0.9525	0.8955	0.9485	0.952	0.9515	1	0.9275	0.9315	0.9385	1
Length	0.1709	0.2724	0.182	0.1805	0.438955	0.1792	0.1416	0.1715	0.1826	0.1813	0.440591	0.1667	0.1664	0.1699	0.43646

Table 5: *Estimated coverage probabilities and widths of the intervals using chi-square distribution* ($M = 2000$).

	Mill^3	S&K	McK	MMcK	C&P	Panich	Prop	Med Mill	Med McK	Med MMcK	MedC&P	NPBS	PBS	BSMill	BS C&P
N=15, CV=.1															
Coverage	0.9195	0.208	0.958	0.9585	0.913	0.9585	0.9565	0.932	0.9505	0.9505	0.9245	0.827	0.8625	0.909	0.902
Length	0.0735	0.0101	0.0844	0.0842	0.071006	0.0812	0.083	0.0749	0.086	0.0858	0.072382	0.0626	0.0626	0.0725	0.070002
N=15, CV=.3															
Coverage	0.9175	0.564	0.9605	0.9615	0.912	0.966	0.946	0.9275	0.952	0.953	0.9225	0.8335	0.867	0.9055	0.901
Length	0.2406	0.0954	0.2969	0.2875	0.232453	0.2752	0.2509	0.2472	0.3038	0.2964	0.238825	0.1917	0.1917	0.2376	0.229511
N=15, CV=.5															
Coverage	0.9275	0.777	0.9805	0.98	0.921	0.9745	0.9405	0.9415	0.971	0.978	0.938	0.8415	0.8675	0.9235	0.9165
Length	0.4473	0.2713	0.7173	0.6183	0.43216	0.5775	0.4127	0.465	0.7387	0.6593	0.449276	0.32	0.3195	0.4401	0.425139
N=25, CV=.1															
Coverage	0.925	0.2825	0.9395	0.9395	0.957	0.944	0.9365	0.927	0.9375	0.9385	0.9605	0.868	0.887	0.9185	0.954
Length	0.0565	0.0101	0.0611	0.061	0.071407	0.0598	0.0603	0.0571	0.0618	0.0617	0.072231	0.0509	0.0508	0.0559	0.070711
N=25, CV=.3															
Coverage	0.9305	0.694	0.9645	0.9635	0.966	0.961	0.946	0.9365	0.957	0.958	0.971	0.865	0.8885	0.929	0.9635
Length	0.1833	0.0938	0.2057	0.2027	0.231837	0.1978	0.1814	0.1866	0.2089	0.2066	0.235998	0.153	0.1532	0.1818	0.229914
N=25, CV=.5															
Coverage	0.949	0.884	0.9705	0.9715	0.9725	0.97	0.93	0.9515	0.957	0.9615	0.9745	0.8755	0.8935	0.9485	0.9695
Length	0.3446	0.2724	0.4318	0.4112	0.435865	0.3978	0.3014	0.3556	0.442	0.4286	0.449855	0.26	0.2603	0.3409	0.431175
N=50, CV=.1															
Coverage	0.953	0.3895	0.961	0.961	0.997	0.9615	0.9585	0.956	0.9575	0.9575	0.9975	0.915	0.9295	0.9495	0.9965
Length	0.0398	0.0101	0.0413	0.0413	0.071908	0.0409	0.0409	0.04	0.0416	0.0416	0.072358	0.0373	0.0374	0.0395	0.071459
N=50, CV=.3															
Coverage	0.9475	0.8605	0.9585	0.9595	0.998	0.9615	0.9445	0.95	0.953	0.9535	0.999	0.9135	0.917	0.9455	0.9975
Length	0.1293	0.0943	0.1368	0.136	0.233722	0.1343	0.1232	0.1309	0.1382	0.1376	0.236501	0.113	0.1131	0.1284	0.232103
N=50, CV=.5															
Coverage	0.9645	0.9735	0.979	0.9785	0.998	0.978	0.929	0.9685	0.9645	0.97	0.998	0.9035	0.914	0.9645	0.9985
Length	0.2409	0.2696	0.2687	0.2638	0.435342	0.2599	0.2035	0.2466	0.2734	0.2709	0.445631	0.1916	0.1916	0.2389	0.431806
N=100, CV=.1															
Coverage	0.9475	0.5185	0.959	0.959	1	0.9555	0.9565	0.949	0.9575	0.9575	1	0.9295	0.932	0.948	1
Length	0.028	0.0101	0.0286	0.0286	0.072027	0.0284	0.0283	0.0281	0.0287	0.0287	0.072271	0.027	0.0269	0.0279	0.071638
N=100, CV=.3															
Coverage	0.958	0.9605	0.963	0.963	1	0.963	0.95	0.9565	0.9595	0.9595	1	0.922	0.932	0.9575	1
Length	0.091	0.094	0.0937	0.0934	0.233826	0.0929	0.0852	0.0918	0.0944	0.0943	0.235882	0.0821	0.0821	0.0906	0.232691
N=100, CV=.5															
Coverage	0.954	0.9965	0.969	0.968	1	0.964	0.917	0.9545	0.96	0.9635	1	0.9055	0.9095	0.9505	1
Length	0.1697	0.2692	0.1806	0.1791	0.435943	0.1778	0.141	0.1732	0.1833	0.1831	0.444846	0.142	0.142	0.1683	0.432382

Table 6: *Estimated coverage probabilities and widths of the intervals using gamma distribution* $(M = 2000)$.

	Mill[4]	S&K	McK	MMcK	C&P	Panich	Prop	Med Mill	Med McK	Med MMcK	MedC&P	NPBS	PBS	BSMill	BS C&P
N=15, CV=.1															
Coverage	0.914	0.218	0.948	0.948	0.9075	0.9515	0.947	0.9265	0.9445	0.9455	0.9175	0.8335	0.8595	0.9	0.8955
Length	0.0735	0.0101	0.0844	0.0842	0.071036	0.0813	0.083	0.0749	0.086	0.0858	0.072374	0.0638	0.0638	0.0725	0.070013
N=15, CV=.3															
Coverage	0.913	0.5505	0.967	0.966	0.905	0.963	0.953	0.9255	0.9605	0.9615	0.9205	0.814	0.8485	0.8995	0.894
Length	0.2368	0.0927	0.2913	0.2825	0.228816	0.2704	0.2476	0.2429	0.2977	0.2906	0.234679	0.1887	0.1888	0.2336	0.225688
N=15, CV=.5															
Coverage	0.926	0.7725	0.985	0.983	0.918	0.981	0.951	0.9405	0.978	0.98	0.9325	0.8355	0.866	0.917	0.9135
Length	0.4398	0.2639	0.6845	0.6029	0.424926	0.5634	0.4077	0.4569	0.7043	0.6366	0.441444	0.3156	0.3161	0.4348	0.420057
N=25, CV=.1															
Coverage	0.928	0.269	0.949	0.949	0.965	0.946	0.948	0.9335	0.944	0.9445	0.9695	0.8715	0.8895	0.9215	0.9615
Length	0.0568	0.0102	0.0615	0.0614	0.071874	0.0602	0.0607	0.0575	0.0622	0.0622	0.072733	0.051	0.051	0.0563	0.071211
N=25, CV=.3															
Coverage	0.937	0.6925	0.968	0.968	0.9695	0.963	0.947	0.9415	0.96	0.96	0.9735	0.8755	0.8975	0.9305	0.965
Length	0.1818	0.0924	0.2039	0.201	0.229989	0.1961	0.1802	0.1851	0.2071	0.2049	0.234137	0.1525	0.1528	0.1803	0.228007
N=25, CV=.5															
Coverage	0.944	0.8695	0.9735	0.973	0.9715	0.97	0.9295	0.951	0.9605	0.9655	0.9765	0.863	0.8825	0.9385	0.969
Length	0.3443	0.2722	0.4314	0.4109	0.435489	0.3975	0.3012	0.3551	0.4414	0.4275	0.449127	0.2578	0.2579	0.3406	0.430841
N=50, CV=.1															
Coverage	0.9425	0.397	0.9565	0.9565	0.995	0.955	0.954	0.9465	0.9535	0.953	0.9955	0.919	0.9275	0.9425	0.994
Length	0.0398	0.0102	0.0414	0.0413	0.071965	0.0409	0.0409	0.0401	0.0416	0.0416	0.072405	0.0373	0.0373	0.0396	0.071531
N=50, CV=.3															
Coverage	0.95	0.8565	0.9605	0.961	0.9975	0.964	0.946	0.949	0.954	0.9555	0.9975	0.906	0.9175	0.9445	0.9975
Length	0.1282	0.0929	0.1356	0.1347	0.231671	0.1331	0.1223	0.1296	0.1368	0.1363	0.234208	0.1124	0.1124	0.1272	0.229973
N=50, CV=.5															
Coverage	0.9575	0.9685	0.9765	0.975	0.997	0.9755	0.933	0.9605	0.9625	0.966	0.997	0.898	0.9075	0.9555	0.997
Length	0.2413	0.2704	0.2692	0.2643	0.43611	0.2604	0.2038	0.247	0.2739	0.2715	0.446485	0.1937	0.1934	0.2386	0.431276
N=100, CV=.1															
Coverage	0.9515	0.533	0.953	0.953	1	0.9555	0.95	0.954	0.951	0.9515	1	0.9295	0.9345	0.9525	1
Length	0.0281	0.0102	0.0287	0.0287	0.072294	0.0285	0.0284	0.0282	0.0288	0.0288	0.072542	0.0271	0.0271	0.028	0.071857
N=100, CV=.3															
Coverage	0.954	0.9535	0.958	0.959	1	0.9575	0.939	0.9525	0.9515	0.9525	1	0.918	0.9265	0.949	1
Length	0.0905	0.093	0.0931	0.0929	0.23239	0.0923	0.0847	0.0912	0.0938	0.0937	0.234425	0.0826	0.0828	0.09	0.231312
N=100, CV=.5															
Coverage	0.969	0.996	0.973	0.973	1	0.9745	0.9285	0.9675	0.959	0.963	1	0.92	0.923	0.9645	1
Length	0.17	0.2698	0.1809	0.1794	0.436682	0.1782	0.1412	0.1736	0.1838	0.1836	0.446032	0.1417	0.1418	0.1689	0.433915

Table 7: *Estimated coverage probabilities and widths of the intervals using normal, chi-square and gamma distributions* $(n = 500, M = 2000)$.

	Mill[5]	S&K	McK	MMcK	C&P	Panich	Prop	Med Mill	Med McK	Med MMcK	MedC&P	NPBS	PBS	BSMill	BS C&P
Normal Distribution															
CV=.1															
Coverage	0.9475	0.9475	0.946	0.9455	0.2167	0.9455	0.9465	0.949	0.949	0.9445	0.9445	0.9415	0.942	0.9435	0.9435
Length	0.0126	0.0126	0.0126	0.0125	0.0018	0.01252	0.0125	0.0126	0.0126	0.0125	0.01253	0.0124	0.0124	0.0125	0.01248
N=15, CV=.3															
Coverage	0.947	0.9465	0.9475	0.9455	0.5645	0.9455	0.925	0.9435	0.9435	0.945	0.9445	0.943	0.9405	0.942	0.9415
Length	0.0408	0.0408	0.0407	0.0405	0.0158	0.04041	0.0374	0.0408	0.0408	0.0405	0.04044	0.04	0.04	0.0403	0.04027
N=15, CV=.5															
Coverage	0.95	0.9505	0.952	0.953	0.671	0.9525	0.909	0.949	0.95	0.9535	0.953	0.944	0.9505	0.9485	0.9485
Length	0.058	0.0579	0.0579	0.0571	0.0281	0.05701	0.0498	0.058	0.058	0.0571	0.05705	0.0565	0.0565	0.0568	0.05678
Chi-Square Distribution															
N=25, CV=.1															
Coverage	0.9485	0.9485	0.9485	0.95	0.2245	0.95	0.9455	0.9475	0.9475	0.947	0.947	0.941	0.941	0.9435	0.9435
Length	0.0126	0.0126	0.0126	0.0125	0.0018	0.01252	0.0124	0.0126	0.0126	0.0125	0.01253	0.0123	0.0123	0.0125	0.01246
N=25, CV=.3															
Coverage	0.961	0.962	0.963	0.958	0.566	0.958	0.937	0.947	0.9475	0.951	0.9505	0.943	0.942	0.956	0.956
Length	0.041	0.041	0.0409	0.0407	0.0159	0.04061	0.0375	0.0412	0.0412	0.0409	0.04087	0.0383	0.0383	0.0405	0.04048
N=25, CV=.5															
Coverage	0.9655	0.9655	0.965	0.962	0.7075	0.962	0.935	0.949	0.95	0.953	0.953	0.9415	0.9435	0.962	0.962
Length	0.0578	0.0578	0.0577	0.0569	0.028	0.05688	0.0497	0.0584	0.0585	0.0576	0.05753	0.0523	0.0523	0.0567	0.05663
Gamma Distribution															
N=50, CV=.1															
Coverage	0.9455	0.9455	0.946	0.945	0.2095	0.9445	0.9415	0.944	0.944	0.9445	0.9445	0.9355	0.937	0.942	0.942
Length	0.0126	0.0126	0.0126	0.0125	0.0018	0.01253	0.0125	0.0126	0.0126	0.0126	0.01254	0.0123	0.0123	0.0125	0.01247
N=50, CV=.3															
Coverage	0.9515	0.951	0.95	0.95	0.579	0.9495	0.9325	0.946	0.946	0.9455	0.9455	0.938	0.9345	0.9455	0.9455
Length	0.0408	0.0408	0.0407	0.0404	0.0158	0.0404	0.0374	0.041	0.041	0.0407	0.04066	0.0381	0.0381	0.0403	0.04028
N=50, CV=.5															
Coverage	0.9735	0.974	0.975	0.968	0.7785	0.968	0.9255	0.948	0.9505	0.957	0.957	0.938	0.9415	0.9655	0.9645
Length	0.0781	0.078	0.0779	0.0759	0.0439	0.07585	0.0622	0.0792	0.0795	0.0773	0.07726	0.0673	0.0673	0.0756	0.0755

References

Banik, S. and Kibria, B. M. G. (2010). Comparison of some parametric and nonparametric type one sample confidence intervals for estimating the mean of a positively skewed distribution. *Communications in Statistics-Simulation and Computation*, 39, 361–389.

Banik, S. and Kibria, B. M. G. (2011). Estimating the population coefficient of variation by confidence intervals . *Communications in Statistics-Simulation and Computation*, 40, 1236–1261.

Bernatm, D. H., Lazovich, D., Forster, J. L., Oakes, J. M. and Chen, V. (2009). Area-level variation in adolescent smoking. *Preventing Chronic Disease*, 6(2), 1–8.

Billings, J., Zeitel, L., Lukomink, J., Carey, T.S., Blank, A. E. and Newman, L. (1993). Impact of socioeconomic status of hospital use in New York City. *Health Affairs*, 162–173.

Center for Disease Control and Prevention (2009). Health, United States, 2009. Retrieved from: http://www.cdc.gov/nchs/data/hus/hus09.pdf.

Curto, J. D. and Pinto, J. C. (2009). The coefficient of variation asymptotic distribution in the case of non-iid random variables. *Journal of Applied Statistics*, 36(1), 21–32.

D'Agostino, Ralph B. and Stephens, Michael A. (1986). Goodness-of-fit-techniques (Statistics: a Series of Textbooks and Monographs, vol. 68).

Efron, B. (1979). Bootstrap methods: another look at the jack knife. *Annals of Statistics*, 7, 1–26.

Efron, B. (1987). Better bootstrapping confidence intervals. *Journal of American Statistical Association*, 82(397), 171–185.

Faber, D. S. and Korn, H. (1991). Applicability of the coefficient of variation for analyzing synaptic plasticity. *Biophysical Society*, 60, 1288–1294.

Hosmer D. W. and Lemeshow, S. (2000). *Applied Logistic Regression*: Second edition. New York: John Wiley & Sons.

Kelly, K. (2007). Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach. *Behavior Research Methods*, 39(4), 755–766.

Kibria, B. M. G. (2006). Modified confidence intervals for the mean of the Asymmetric distribution. *Pakistan Journal of Statistics*, 22(2), 109–120.

Koopmans, L. H., Owen, D. B. and Rosenblatt, J. I. (1964). Confidence intervals for the coefficient of variation for the normal and log normal distributions. *Biometrika Trust*, 51(1/2), 25–32.

McGibney, G. and Smith, M. R. (1993). An unbiased signal?to?noise ratio measure for magnetic resonance images. *Medical Physics*, 20(4), 1077–1079.

McKay, A. T. (1932). Distribution of the coefficient of variation and the extended t distribution. *Journal of Royal Statistical Society*, 95, 695–698.

Miller, E. G. (1991). Asymptotic test statistics for coefficient of variation. *Communication in Statistic-Theory & Methods*, 20, 3351–3363.

Panichkitkosolkul, W. (2009). Improved confidence intervals for a coefficient of variation of a normal distribution. *Thailand Statistician*, 7(2), 193–199.

R Development Core Team (2010). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL http://www.R-project.org.

SAS Institute Inc. (2008). SAS/STAT User's Guide, Version 9.2, Cary, NC: SAS Institute Inc.

Sharma, K. K. and Krishna, H. (1994). Asymptotic sampling distribution of inverse coefficient-of-variation and its applications. *IEEE Transactions on Reliability*, 43(4), 630–633.

Shi, W. and Kibria, B. M. G. (2007). On some confidence intervals for estimating the mean of a skewed population. *International Journal of Mathematical Education in Science and Technology*, 38(3), 412–421.

SPSS for Windows (2008). SPSS Inc., Chicago, IL.

Vangel, M. G. (1996). Confidence intervals for a normal coefficient of variation. *The American Statistician*, 15(1), 21–26.

Visintainer P. F. and Tejani, N. (1998). Understanding and using confidence intervals in clinical research. *The Journal of Maternal-Fetal Medicine*, 7, 201–2.