MJFAS
Malaysian Journal of Fundamental and Applied Sciences

REVIEW ARTICLE

Confidence Intervals by Bootstrapping Approach: A Significance Review

Siti Fairus Mokhtar[a,b]*, Zahayu Md Yusof[a,c], Hasimah Sapiri[a]

[a]School of Quantitative Sciences, Universiti Utara Malaysia 06010 Sintok, Kedah, Malaysia; [b]Mathematical Sciences Studies, College of Computing, Informatics and Media, Universiti Teknologi MARA (UiTM) Kedah Branch, Sungai Petani Campus, Malaysia; [c]Institute of Strategic Industrial Decision Modelling (ISIDM), Universiti Utara Malaysia, Kedah, Malaysia

Abstract A confidence interval is an interval estimate of a parameter of a population calculated from a sample drawn from the population. Bootstrapping method, which involves producing several new data sets that are resampled from the original data in order to estimate parameter for each newly created data set, allowing an empirical distribution for the parameter to be estimated. Since certain statistics are harder to estimate, confidence intervals are rarely employed. Several statistics might necessitate multi-step formulas assuming that are impractical for calculating confidence intervals. This paper reviews research on the concept of bootstrapping and bootstrap confidence interval. The current narrative analysis was developed to answer the main research question: (1) What is the concept of the bootstrap method and bootstrap confidence interval? (2) What are the methods of bootstrapping to obtain confidence interval? This study has found general bootstrap method idea, various techniques of bootstrap methods, its advantages and disadvantages, and its limitations. There are normal interval method, percentile bootstrap method, basic method, first-order normal approximation method, bias-corrected bootstrap, accelerated bias-corrected bootstrap and bootstrap-*t* method. This study concludes that the advantages of using bootstrap CI is that it does not require any assumptions about the shape of distribution and universality of the approach. Bootstrapping is a computer-intensive statistical technique that relies significantly on modern high-speed digital computers to do massive computations.

Keywords: Bootstrap, confidence interval, parameter estimations, resampling.

***For correspondence:**
fairusmokhtar706@gmail.com

Received: 27 June. 2022
Accepted: 28 Dec. 2022

Introduction

In statistics, the information in sample $X = (X_1, X_2, \ldots, X_n)$ is used to estimate an unknown parameter (Salkind 2010). There are several methods in estimating a value of a parameter, such as by using hypothesis testing, point estimation and confidence interval (CI) estimation. A CI is an interval estimate of a parameter of a population calculated from a sample drawn from the population. An unknown population parameter's interval estimate generated from a sample chosen from that population. Whenever there is a known and statistically reasonable level of confidence that the unknown population parameter resides within that interval, it is described as CI (Petty 2012). CI is an important statistical estimator of population location and dispersion parameters (Abu-Shawiesh, Sinsomboonthong, and Kibria 2022). The CI is the upper bound and lower bound that contains a true mean value. CI gives the probability of the population parameter of interest (Tapia, Salvador, and Rodríguez 2020). A confidence level is utilized to construct interval estimates. This is the probability that the interval corresponds to the population parameter that is being estimated. Despite 95% confidence level which certainly the most commonly used, confidence intervals can be calculated at any confidence level, for instance, 90% or 99%.

CI provides additional information than point estimates (Das, 2019). A point estimate is a single value (i.e., mean, median) that is an unknown parameter based on statistics. CI is a range calculated from observable data that includes the true value of an unknown population. Due to the procedure's tendency to generate intervals that contain the population parameter, CI is a good estimate of the parameter (Das, 2019). CI is wonderful basis for inference because give point estimate together (Cumming, 2007). CIs are consisting of a margin of error and point estimate around that point estimate. On the other hand, a parameter's of CI estimation is a range of probable values. Therefore, the CI estimation is more suitable than point estimation because it gives statistically significant information (Saito & Dohi, 2018). Figure 1 shows a type of estimation.

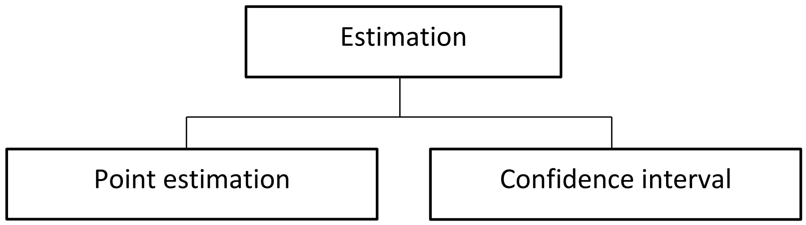

Figure 1. Type of estimation (Das, 2019)

The researcher tend take a small sample size when time constraints and costs are high. Then, estimation for the population parameters will be more suitable if it can be generated using statistics from the small samples. However, there are three basic assumptions of validity of CIs that are challenging to meet with small sample sizes (LaFontaine 2021). LaFontaine (2021) summarized that once a representative sample is obtained, those inferences can be made. There are three assumptions for the accuracy of these inferences. The assumption of normality of the sampling distribution of the parameter is the first of these three assumptions. The standard assumption of the CI is the residual must be identically and independently distributed with normal distribution. Depending on the standard normal and *t*-distributions, the common CI is symmetric at about zero (Berrar 2019; Dogan 2004). The second assumption is that the estimated parameter's standard error is a close approximation to the standard deviation of the estimated parameter's sampling distribution. The last of the assumptions is the estimated parameter has little bias in its estimate. For some parameters, these assumptions can be met relatively easily, while for other parameters it requires a different set of methods to meet these assumptions. These assumptions need to be fulfilled when making inferences on the population mean. Violation of assumptions will create problems, especially when estimating the CIs.

The uncommon use of CIs is related to estimation difficulties for specific statistics, according to a study by Banjanovic & Osborne (2016). Some statistics may necessitate multi-step formulas with assumptions that are not constantly inconvenient to be used when generating CIs. The traditional CI are computed using the Equation 1 and 2. The 95% CIs are calculated based on formula.

$\bar{x} \pm 1.96\left(s/\sqrt{n}\right)$ for large data (1)

$\bar{x} \pm t_{\alpha/2}\left(s/\sqrt{n}\right)$ for small data (n less than 30), (2)

in which $\bar{x}$ denotes the sample mean, s is the standard deviation, n represents the sample size, $s/\sqrt{n}$ is standard error, and $t_{\alpha/2}\left(s/\sqrt{n}\right)$ is margin.

The traditional approach is accompanied by two problems. First off, the distribution is symmetric and relies on the normality assumption. As a result, the normality assumption also forms the basis of the (1 – α)100% CI with respect to the population mean (μ). Nonetheless, real life does not support the normality assumption. The traditional approach is not particularly reliable in these circumstances by many other researchers, which include (Boos and Hughes-Oliver 2000; David 1998; Desharnais *et al*. 2015; Wilcox 2021). Although its average variability and widths were not as low as other CI, prior studies have showed that the traditional approach fits well for asymmetric distributions and small sample sizes with respect to the coverage probability approaching the nominal confidence coefficient (Boos and Hughes-Oliver 2000; Shi and Golam Kibria 2007; Wang 2001; Zhou and Dinh 2005).

Traditional approach to statistical inference is based on assumptions. The weakness of this traditional approach is that it is not robust under extreme deviations from normality. Normal approximation theory is imprecise for limited data and lack estimate of the precision (Saha and Kapilesh 2016). When the assumption of normality is not followed, the bootstrapping method is an alternative technique to estimate the parameter (L. I. Tong, Saminathan, and Chang 2016; Flowers-Cano *et al*. 2018). Bootstrapping is a method that relaxes normality, independence, and constants variation assumptions (Hongyi Li and Maddala 2007; L. I. Tong, Saminathan, and Chang 2016; Rousselet, Pernet, and Wilcox 2021).

Bootstrapping, jackknife resampling, permutation test, cross-validation, and Monte Carlo simulation are examples of resampling techniques. When sample size and population-representative are limited, the bootstrapping method is used. The sample is assumed to be connected to the population in the same way that an empirical distribution created via resampling *N* samples with substitutions from the original distribution, each of the same sizes as the original sample, is connected to the population. Moreover, the researcher can assess the accuracy of the predictions on the population parameter by producing an empirical distribution and comparing it to the sample statistic (LaFontaine, 2021). The distribution of observed data such as data in a random sample is the empirical distribution. The values of the statistic can generate standard errors and CIs for the statistic (L. I. Tong *et al*., 2016). The use of bootstrapping CI in reporting the results can be seen in the area of biomedical research, financial investment, and off-policy evaluation (Hanna *et al*., 2017; Haukoos & Lewis, 2005; Klaudia & Łukasz, 2020).

The process of generating a pseudo sample from the original sample or applying a model to the original sample is known as the bootstrapping principle (Puth *et al*., 2015). The purpose of this paper is to comprehend CI utilising various bootstrap techniques. This includes first-order normal approximation, percentile bootstrap method, bias-corrected bootstrap, normal interval method, accelerated bias-corrected bootstrap, basic method, as well as bootstrap-*t* method. This article attempts to introduce the reader to the concepts and methods of bootstrap in statistics, which is under a larger umbrella of resampling. Although previous research has study about concept of bootstrapping, which has been a limitation in most studies. The procedure of the bootstrap method has been discussed to provide an understanding of application of bootstrap methods in the calculation of CI.

This paper is organized as follows. The first section provides a summary of the bootstrapping method and the concept of bootstrapping method. The next section discusses different methods of bootstrapping to obtain a CI, process of bootstrapping method, and the performance of CI. The final section concludes about overall bootstrapping method.

Bootstrap Method in General

Efron (1979) proposed the bootstrap method which is a databased simulation method for statistical inferences. Bootstrap method can be applied for interval estimate for mean, standard deviation, or any other statistic with the assumption that the observations are independent and come from the same distribution (Good & Hardin, 2012). Bootstrap method is a method for resampling a model or one's data inferred from the data to estimate the distribution of an estimator or test statistic (Horowitz, 2019).

Given its generality, the bootstrapping method is a common method for generating CIs. A similar procedure may be utilized to a wide variety of statistics (Puth *et al*., 2015). The goal of bootstrapping is to conclude a population parameter based on the data available (Rousselet *et al*., 2021). Bootstrap method is more generalized and versatile than traditional method because bootstrap method is computationally intensive and efficiently applied for uncertainty analysis (Saha & Kapilesh, 2016; L. I. Tong *et al*., 2016). Bootstrapping is used when sample sizes are small, the distribution of estimators is unknown or when the satisfaction of relevant assumptions is not met (Bochniak, Kluza, Kuna-Broniowska, & Koszel, 2019). Bootstrapping method with robust estimators (for example, M-estimators, trimmed means, median) can aid people in comprehending data more deeply than traditional means. Robust estimators do not react very strongly to outliers. Advantageously, the bootstrapping method does not become reliant on underlying population assumptions (L. I. Tong *et al*., 2016). Bootstrap methods are thought to have two primary benefits over more conventional ways: (a) they are more robust, coping effectively with data that does not comply with standard parametric assumptions, and (b) they are frequently easier to enforce, relying on computer power to replace complex derivations (Hoyle & Cameron, 2003). Therefore, this study will explore and understand the concept of bootstrapping.

Generally, bootstrap method also known as percentile bootstrap CI (Pek, Wong, and Wong 2017; Abu-shawiesh and Saeed 2022). Although CI development tends to involve a complicated resampling technique with a good theoretical coverage probability performance, the CI might perform erratically in practice depending on the bootstrap estimator's distribution (Sinsomboonthong, Abu-Shawiesh, and

Kibria 2020). Additionally, this approach is difficult to put into practice since it requires statistical programming to compute, in contrast to bootstrap-*t*, which is suggested in this study and is simple to put into use.

Concept of bootstrap method and bootstrap CI

Suppose $x = x_1, x_2, \ldots, x_n$ is a random sample of size *n* drawn from original sample. Generate a bootstrap sample from the replacement sample. There is a total of n^n resamples feasible. Bootstrap sampling is the same as sampling from the empirical population distribution with replacement. Then, estimate the parameter value for each bootstrap sample.

In the bootstrapping method, the number of replications is important because it acts as a guideline for an efficient and accurate bootstrap resampling (Hedges 1992). According to DiCiccio & Efron (1996) and Yan (2022), at least 2000 replications are used when conducting bootstrapping methods. The number of bootstraps cannot be less than n^n (Efron and Tibshirani 1993). (L. I. Tong, Saminathan, and Chang 2016; Zhao *et al*. 2021; Wilcox 2021) suggest that to get reasonable accuracy in terms of CI estimates for the parameter, at least 1000 bootstrap resamples are adequate. Let *B* be the number of bootstrap samples to be derived from the original dataset. Therefore, a much bigger value of *B* is required for bootstrap CI (Efron and Tibshirani 1993). Banjanovic & Osborne (2016) recommended that more bootstrap samples will improve the estimation and take a modern computer only slightly longer by using 5000 bootstraps resamples (L. I. Tong, Saminathan, and Chang 2016). The bootstrap method was found to be robust to the choice of initial sample data, the number of bootstrap samples, and the different bootstrap samples obtained in different runs. The study suggests that numbers for estimation as 4,000 to 5,000 bootstrap samples, each of 30 or more data (Saha and Kapilesh 2016). Results tend to be more consistent when the number of bootstraps increases (Mahmudah *et al*. 2023). Saha & Kapilesh (2016) also suggests that the number of bootstrap samples between 1,000 and 2,000 bootstrap samples, with each sample containing at least 25 data in each sample. Study by L. I. Tong *et al*. (2016), for small sample size the data used are *n*=5,6,…,30.

Assume that the population parameter θ is estimated using a random sample. The bootstrap estimate of θ is expressed by $\widehat{\theta}^*$. To get B sets of bootstrap samples, the resampling process is performed B times. Note that B should represent how many bootstrap samples will be drawn from the original dataset. The bias-corrected bootstrap, bias-corrected accelerated percentile bootstrap CI, standard bootstrap CI, as well as percentile bootstrap are the four bootstrap CIs (L. I. Tong, Saminathan, and Chang 2016). Figure 2 depicts a schematic representation of steps in bootstrapping (Haukoos and Lewis 2005).

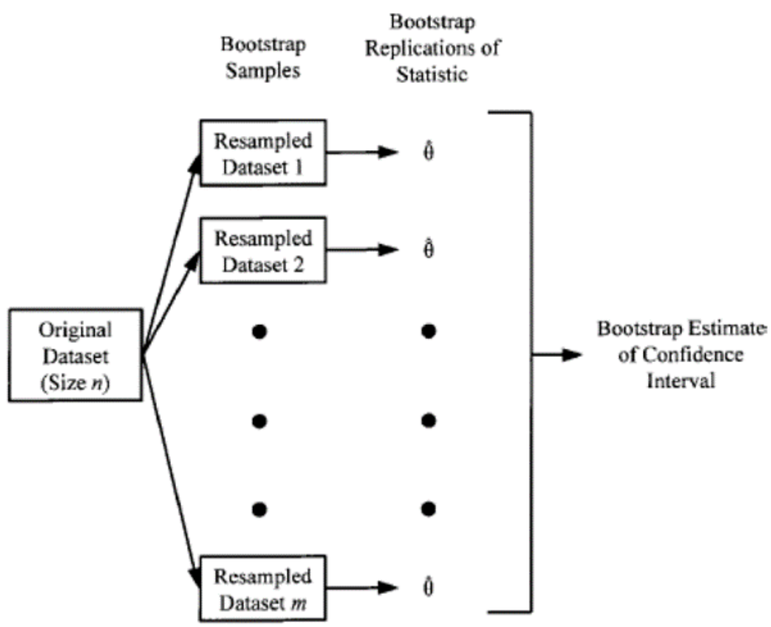

Figure 2. Description of the steps in bootstrapping

Parametric and Nonparametric bootstrap method

A bootstrap method is known in two different forms. Nonparametric bootstrap and parametric bootstrap. Nonparametric bootstrapping refers to bootstrapping from the original sample, as opposed to parametric bootstrapping, which is bootstrapping from the model that was fitted to the original sample (Puth *et al*., 2015). A sample distribution assumption is established before employing the parametric bootstrap method. In a normal distribution, for instance, two parameters are required, whereas, in a Poisson distribution, only one parameter is required. The statistic is estimated by employing sampling with replacement in the nonparametric method, and the statistic's distribution is attempted to be identified. Nonparametric bootstrapping where the parameters are resamples with replacement to create many bootstrap replicate dataset.

The basic idea behind the parametric bootstrap method is to estimate parameters utilizing the original observation data and then sampling after acquiring the appropriate parameter distribution (Fang & Zhang, 2013; Onyesom & S.I., 2021). The parametric bootstrap is a method of sampling from a parametric probability function that was created by fitting residual terms into an idealised probability distribution. Since it does not necessitate a big data set, when the residual term fits a parametric distribution, this approach works well. By sampling error terms from the parametric distribution after the distribution parameters have been established, a new data set can be created. After the distribution parameters have been determined, the distribution can be used to construct a new data set by sampling error terms from the parametric distribution. In parametric bootstrap, the corresponding probability distribution parameters would be used for resampling process (Saha & Kapilesh, 2016). Nevertheless, with this procedure, distribution assumptions should be fulfilled (Mesabbah *et al*., 2015).

On the contrary, the nonparametric bootstrap method works by generating many fresh bootstrap samples and resampling the statistics (Fang & Zhang, 2013). Nonparametric sampling is the basic sampling technique since it does not involve any distributional assumptions (Puth, Neuhäuser, and Ruxton 2015); However, it does necessitate a large data sample to be appropriately relevant. Nonparametric bootstrap method suited for limited data sets (Saha and Kapilesh 2016). In this technique, the set of residuals ($\hat{\varepsilon}_{it}$) corresponding to the best-fit parameters (θ') or the original parameter values ($\tilde{\theta}$) is resampled at random with replacement to produce each new set of residuals. In the parametric method, no corresponding assumption is required (Mesabbah, Rashwan, and Arisha 2015) while in nonparametric bootstrap, equal probability is assigned to each observation (Saha and Kapilesh 2016). Table 1 compares two different types of bootstrap method (parametric and non-parametric) with traditional counterpart.

Table 1. Comparison two different types of bootstrap method (parametric and non-parametric) with traditional

	Nonparametric bootstrap	Parametric bootstrap	Parametric (traditional)
Assumption	Large sample size	Know the distribution	Need to fulfil assumptions
Use	Traditional formulas not available or too difficult	Traditional formulas not available or too difficult	Have population data and able to use traditional formulas
Sample	The observations with replacement	Estimate from true distribution	True population
Advantages	Can be used to estimate any sampling distribution	More powerful than nonparametric bootstrap	Formulas are wide use
Disadvantages	Not suitable for small sample sizes	Can be difficult to choose distribution	Can only estimate sampling distributions under tight conditions and assumptions

Bootstrap Method to Obtain a CI

Several methods to calculate CI based on bootstrap method require different assumptions. There is a different method of obtaining a CI using bootstrapping. There are normal interval method, percentile bootstrap method, basic method, first-order normal approximation method, bias-corrected bootstrap, accelerated bias-corrected bootstrap and bootstrap-*t*. The package that suggests for simulation is Scilab, R package, and STATA.

(i) Basic Method

Puth *et al*., (2015) indicated that each $\hat{\theta}_i^*$, an error e_i is calculated as $\hat{\theta}_i^* - \hat{\theta}$. It is these e_i that are then ordered, and identify the lower and upper limits e_l and e_u that enclose the central $100(1-\alpha)\%$ CI for the population parameter is then $\left[\hat{\theta} - e_l, \hat{\theta} - e_u\right]$.

(ii) Normal Interval method

The bootstrap distribution is used in the normal interval approach to estimate the standard error for calculating the traditional CI. The CI is computed by using $\theta^* \pm 1.96 * SE$,where $\theta^*$is the sample estimate, as well as SE, which is known as standard error (Banjanovic and Osborne 2016).

(iii) Percentile bootstrap method

The percentile bootstrap is the most basic type of bootstrapping, with bootstrap samples sorted from smallest to biggest. According to Mesabbah *et al*. (2015), the quantile of the bootstrap distribution of the parameters estimate is utilized in the percentile bootstrap CI technique. Efron proposed the percentile bootstrap CI in 1982. The percentile bootstrap interval is simply the interval between the $100\left(\alpha/2\right)$ and $100\left(1-\alpha/2\right)$ percentiles of the distribution of *θ* estimates obtained from resampling, where *θ* denotes a parameter of interest and *α* is the significance level. Furthermore, the upper and lower bounds are the exact percentiles corresponding to the given alpha level. For example, *α* = 0.05 for 95% CIs. The $\alpha/2$ and $1-\alpha/2$ quantiles of the bootstrap distribution are utilized to construct the CI bounds, .025 and .975 respectively.

The following is the method to generate a percentile bootstrap CIs of $\hat{\theta}$ (an estimator of θ):
(1) *B* random bootstrap samples are created,
(2) each bootstrap sample yields a parameter estimate,
(3) all *B* bootstrap parameter estimates are sorted from lowest to highest, and
(4) the CI is [$\hat{\theta}$lower limit, $\hat{\theta}$upper limit] =[$\hat{\theta}_j^*$, $\hat{\theta}_k^*$], in which $\hat{\theta}_j^*$ expresses the j^{th} quantile (lower limit), as well as $\hat{\theta}_k^*$ expresses the k^{th} quantile (upper limit); j=[$\frac{\alpha}{2}$ ×*B*], k=[(1− $\frac{\alpha}{2}$)×*B*].
For instance, a 95% percentile bootstrap CI having 1,000 bootstrap samples is the interval between the 975^{th} quantile value and the 25^{th} quantile value of the 1,000 bootstrap parameter estimates (Jung *et al*. 2019). However, CI is often biased when small in size (Shao and Tu 1995). Percentile bootstrap tend to be inaccurate because the bootstrapping for the sampling distribution is skewed distribution and bias (Rousselet, Pernet, and Wilcox 2021).

(iv) First-order normal approximation

Puth *et al*., (2015) identified that this method, if $\hat{\sigma}_B$ is the standard deviation of the *B* bootstrap samples and $\bar{\theta}$ is the mean, then the $100(1-\alpha)\%$ CI is given by $\bar{\theta} \pm z_{\alpha/2}\hat{\sigma}_B$ where $z_{\alpha/2}$ is the *z*-score for a given level of significance α, if $\alpha = 0.05$ then $z_{\alpha/2} = 1.96$.

(v) Bias-Corrected Bootstrap

The bootstrap empirical distribution may occasionally be asymmetrical distribution, resulting in a bias in the confidence interval (L. I. Tong, Saminathan, and Chang 2016). The percentile bootstrap confidence interval is thus corrected (Efron 1981).

First, P_0 is computed using the bootstrap distribution of $\hat{\theta}^*(i)\ (i = 1, 2, \ldots, B)$ as follow:
(1) $P_0 = Pr\left(\hat{\theta}^*(i) \le \theta\right),\ (i = 1, 2, \ldots, B)$
(2)Next, z_0 is computed to measure the bias of bootstrap distribution as follows:

$z_0 = \theta^{-1}(P_0)$ where $\theta^{-1}(.)$ is the inverse cumulative standard normal distribution.
(3)The bias-corrected percentiles P_L and P_U are computed, respectively, as given in the formulas:

$$P_L = \phi\left(2\hat{z}_o - z_{1-\alpha/2}\right)$$

$$P_U = \phi\left(2\hat{z}_o + z_{1-\alpha/2}\right)$$

where $\phi(.)$ is the cumulative standard normal distribution. Thus, the $(1-\alpha)\,100\%$ bias-corrected bootstrap confidence interval for θ can be computed as $\left(\hat{\theta}^*(P_L B), \hat{\theta}^*(P_U B)\right)$.

Although these intervals may never be like the percentile method, the bias-corrected bootstrap employs percentiles as the upper and lower CIs. The bias-corrected bootstrap is employed to recalculate the endpoints of the CI in place of the percentile method. According to Chen & Fritz (2021), there are six different bias corrections: (a) mean, (b–e) Winsorized mean with 10%, 20%, 30%, and 40% trimming in each tail, as well as (f) medcouple (robust skewness measure). Therefore, if no bias is present, the bias-corrected bootstrap will be equivalent (Chen and Fritz 2021).

(vi) Accelerated Bias-Corrected Bootstrap
The accelerated bias-corrected bootstrap (BCa) approach alters both skewness and bias of the bootstrap parameter estimates by incorporating a bias-correction factor as well as an acceleration factor (Efron 1987; Efron and Tibshirani 1993). The BCa formulae may be found in the equation given as follows.

The bias-correction factor $\hat{z}_0$ is calculated as the percentage of bootstrap estimates that are smaller than the initial parameter estimate $\hat{\theta}$,

$$\hat{z}_0 = \phi^{-1}\left(\frac{\#\{\hat{\theta}^{\cdot} < \hat{\theta}\}}{B}\right)$$

where ϕ^{-1} is the inverse function of a standard normal cumulative distribution function (for example, ϕ^{-1} (0.975) = 1.96). Therefore, the bias correction bootstrap percentile CI is as follows: $\left[\hat{\theta}^{\cdot\alpha_1}, \hat{\theta}^{\cdot\alpha_2}\right]$ in which α_1 and α_2 are modified quantities of the location of the CI's endpoints. The endpoints of the CIs at a significant level of 100α% are described as follows: $\alpha_1 = \phi\left(2\hat{z}_0 + Z^{\alpha/2}\right)$, and $\alpha_2 = \phi\left(2\hat{z}_0 + Z^{1-\alpha/2}\right)$ where is the cumulative standard normal distribution (Mesabbah, Rashwan, and Arisha 2015).

The drawback of the BCa method is that it is difficult to estimate the parameter because of the complicated formula (Hoyle and Cameron 2003). In this scenario, the jackknifing method is used to estimate parameters (Hoyle and Cameron 2003). BCa method can be insufficient for a small sample size (Good 2005).

(vii) Bootstrap-*t*
A bootstrap-*t* method is termed as the studentised bootstrap or percentile-*t* bootstrap. Efron and Tibshirani suggested the bootstrap-*t* method in 1993. For various sample sizes, the bootstrap-*t* method surpasses the Student's *t*-test (Zhao *et al*. 2021). The construction of the bootstrap-*t* is similar to the constructions of CIs for the expected value based on random variables with normal distribution. The percentile bootstrap is recommended for inferences about a 20% trimmed mean, although the bootstrap-*t* can yield more exact CIs for the mean, including certain trimmed means (Rousselet, Pernet, and Wilcox 2021). Bootstrap-*t* method constructs CIs that are more accurate and less biased than other bootstrapping methods since the skewness in distribution was handled poorly by all bootstrapping methods except bootstrap-*t* (Hoyle and Cameron 2003).

Bootstrap-*t* method can be used to estimate the distribution of statistic $Z = \frac{\hat{\theta} - \theta}{\widehat{se}}$ directly from the data. For each set of bootstrap samples, a value of Z can be estimated as follows.

$$Z^*(b) = \frac{\hat{\theta}^*(b) - \theta}{\widehat{se}^*(b)}$$

in which $\hat{\theta}^*(b)$ resambles the estimate of $\hat{\theta}$ for the b^{th} bootstrap sample as well as $\widehat{se}^*(b)$ represents the estimated standard error of $\hat{\theta}^*$ for the b^{th} bootstrap sample. The α^{th} percentile of $Z^*(b)$ is estimated as the value $\hat{t}^{(\alpha)}$. According to Barker (2005), the CI is then calculated by using $\left(\hat{\theta} - \hat{t}^{(1-\alpha)} \times \widehat{se}, \hat{\theta} - \hat{t}^{(\alpha)} \times \widehat{se}\right)$
Table 2 summarizes seven methods of bootstrap methods.

Table 2. Summary table for 95% CI estimation using bootstrap methods (Banjanovic & Osborne; 2016)

Method	Formulae	Description
Normal Interval	The CI is computed by $\theta^* \pm z^*_{\alpha/2} SE$ where θ^* is the sample estimate.	- The bootstrap statistic's distribution is approximately symmetrical and normal. - An unbiased estimator of the population estimate is the sample estimate.
Percentile bootstrap	To compute CI, the estimates at the .025 and .975 quantiles of the bootstrap distribution were employed.	- The bootstrap statistic's distribution is essentially normal. - Unbiased estimation of the population estimate is provided by the sample estimate.
Basic	The CI is calculated using $100(1-\alpha)\%$ for the population parameter then $\left[\hat{\theta}-e_l,\ \hat{\theta}-e_u\right]$.	- Based on the premise that the bootstrap distribution of errors is a decent approximation of the actual distribution of sampling errors, this approach is then employed.
First-order normal approximation	The CI is computed via $\bar{\theta} \pm z_{\alpha/2}\hat{\sigma}_B$, in which $z_{\alpha/2}$ denotes the *z*-score for a given level of significance α, if $\alpha = 0.05$ then $z_{\alpha/2} = 1.96$.	- This technique makes use of the finding that bootstraps frequently resemble a normal distribution.
Bias-Corrected Bootstrap	For a 95% CI, the bias-corrected bootstrap adjusts the percentile interval for bias by modifying the percentile points to values other than the 25th and the 975th quantile values of the 1000 bootstrap (Shao and Tu 1995).	-This method assumes that data transformations can achieve normality and a constant standard error (Efron 1987).
Accelerated Bias-Corrected Bootstrap	The CI is determined at the .025 and .975 quantiles of the corrected distribution after bias (skew) and acceleration (nonconstant variance) are corrected from the bootstrap distribution.	-BCa approach alters both skewness and bias of the bootstrap parameter estimates by adding an acceleration factor in addition to a bias-correction factor.
Bootstrap-*t*	The CI is calculated using $\left(\hat{\theta}-\hat{t}^{(1-\alpha)}\times\widehat{se}, \hat{\theta}-\hat{t}^{(\alpha)}\times\widehat{se}\right)$	- Bootstrap-*t* method also known as percentile-*t*.

Process of bootstrapping

In bootstrap method, the bootstrap sample is of same size *n* as original sample. Observations could be repeated in bootstrap sample. The procedure is further repeated, *B* times. A general bootstrapping step is as follow (L. I. Tong, Saminathan, and Chang 2016).

(i) Generate a random sample. This dataset is called the original sample $x = x_1, x_2, \ldots, x_n$.

(ii) Draw *n* samples, with replacement randomly from the original sample to form one bootstrap sample $x^*_1, x^*_2, x^*_3, \ldots, x^*_n$.

(iii) Generate bootstrap dataset *B* times (with replacement) from original sample each with size *n* $x^*_{(i)} = \left(x^*_1, x^*_2, \ldots, x^*_n\right), i = 1, 2, \ldots, B$.

(iv) Compute the parameter from *n* sample $(i = 1, 2, \ldots, B)$

(v) Construct CI by using seven bootstrap methods of statistic of interest from the bootstrap samples.

(vi) Repeat (ii)-(v) *N* times to obtain set of bootstrap CI. Then, calculate performance of seven bootstrap CI evaluated using three indices.

The flowchart of the bootstrapping method is presented in Figure 3. The flowchart illustrates the steps to compute the CI using bootstrapping method.

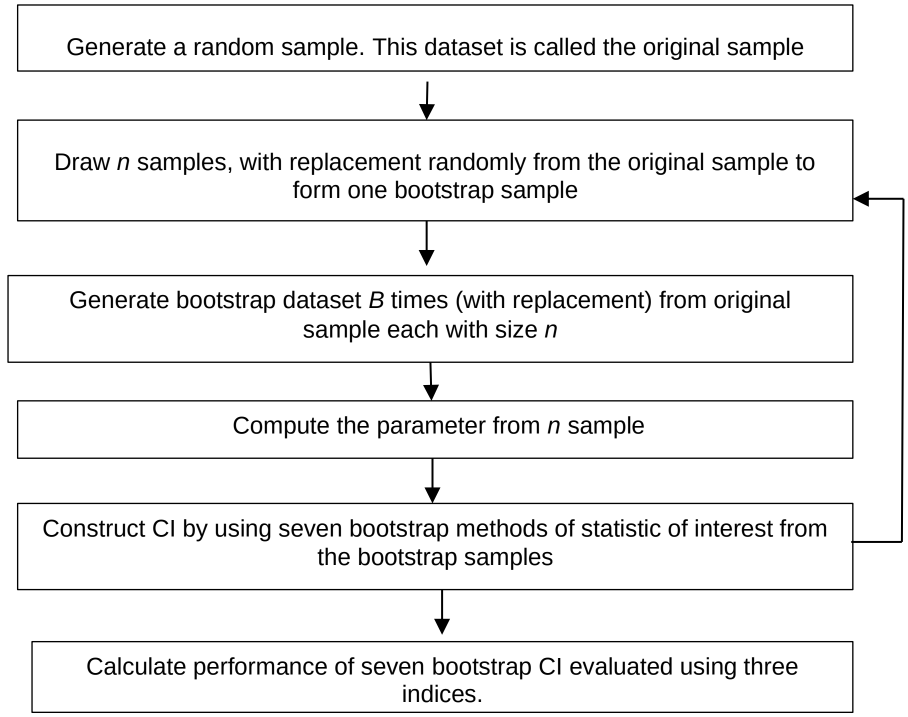

Figure 3*.* Flowchart of bootstrapping method (Saha & Kapilesh, 2016)

Performance of bootstrapping methods

The performance indices are taken after computing the CI estimation. The performance indices were utilized to examine the precision and accuracy of the uncertainty for a variety of estimations. There are three indices that were utilized to assess the performance of bootstrap CI are coverage performance index, interval mean index and interval standard deviation index (Lee Ing Tong *et al*. 2012; Chou *et al*. 2006; L. I. Tong, Saminathan, and Chang 2016). Initially, there is the coverage performance index, which shows what percentage of the time the actual emission fits inside the bootstrap CIs. The bootstrap CI estimate is more accurate when the performance index value is higher. The percentage of times the CIs contain the actual amount of interest is what is meant by coverage. The confidence level of the interval should match this coverage. In reality, coverage is deemed sufficient if the maximum coverage is not significantly higher than 95%, the lowest coverage is not much lower, and the average coverage is close to 95% (Flowers-Cano *et al*. 2018). Second, the interval mean index. The discrepancy between the upper and lower bootstrap CIs is expressed by this index. Accuracy and precision are implied by the smaller interval. Finally, the interval standard deviation index. The standard deviation of bootstrap CI interval lengths is expressed by this index. A small standard deviation indicates a smaller estimated variation and greater bootstrap CI estimates.

Advantages and disadvantages of bootstrapping

There are a few strengths of calculating CI using the bootstrapping method. Firstly, bootstrapping method is use to estimate the CI of uncertainty (e.g., standard deviation) without the assumption of normality of residuals (Endo *et al*., 2015; Klaudia & Łukasz, 2020; L. I. Tong *et al*., 2016). Compared to traditional approaches (i.e., maximum likelihood estimation), the bootstrapping method is more transparent, simpler, versatility approach and general. The advantage of bootstrap method is its simplicity because it is direct in estimation the standard error and CI. Bootstrapping method does not make strong assumptions for the model or the data (Dogan, 2004). The theoretical advantage of the bootstrapping method in calculating CIs is that the population is not necessarily normal distribution (Kennedy & Schumacher, 1993). Secondly, the assumptions of the bootstrapping method are less restrictive and easier to verify compared to the traditional approach (Rousselet *et al*., 2021; L. I. Tong *et al*., 2016). When the traditional approach is difficult or impossible to apply, the bootstrapping method can be used. Employing the bootstrapping method, a statistician may test the statistical accuracy of complex operations by utilizing the computer's capabilities (Barker, 2005). According to Doğan (2017), bootstrapping is an efficient technique for computing CIs. In widely utilized statistical computing software packages, little extra work is required (Puth *et al*., 2015). Finally, bootstrapping methods are flexible and appropriate for a small sample size (Rahmandad *et al.,* 2013; Zhao *et al*., 2021). It however provides accurate results for large samples, regardless of the underlying population (Barker, 2005).

However, disadvantages of bootstrapping method are the need for powerful computers, it must understand the randomness to find the next sample, and large sample sizes must be generated (Dogan, 2004). The sample must represent the values found in the population it was drawn from. If the sample data distribution does not match the population distribution, the bootstrapping method may introduce an additional level of sampling error, resulting in inaccurate statistical estimates. This shows the significance of getting good data that accurately represents the characteristics of the sampled group. The software packages are not easy to compute, and it is time-consuming. Bootstrapping methods also increase the computational costs and implementation challenges (Rahmandad *et al*., 2013). Zaman (2016) in their study described the bootstrap method as a method that includes more computations than typical parametric findings and mathematical analysis and is used as a statistical outcome technique.

Conclusions

This paper reviews research on the concept of bootstrapping and bootstrap CI. The goal of this research is to comprehend CI utilising various bootstrap techniques. The normal interval method, accelerated bias-corrected bootstrap, basic method, first-order normal approximation, bias-corrected bootstrap, percentile bootstrap method, as well as bootstrap-*t* method, were thoroughly discussed in this work. This article attempts to introduce the reader to the concepts and methods of bootstrap in statistics, which is under a larger umbrella of resampling. The procedure of the bootstrap method has been discussed to provide an understanding of application of bootstrap methods in the calculation of CI. Nonparametric methodology like bootstrap better suited than approximation normal (Saha and Kapilesh 2016). The construction of the bootstrap-*t* is similar to the constructions of CIs for the expected value based on random variables with normal distribution. For conclusions regarding a mean that has been reduced by 20%, the percentile bootstrap is advised, although the bootstrap-*t* can yield more exact CIs for the mean, including certain trimmed means (Rousselet, Pernet, and Wilcox 2021). Bootstrap-*t* method constructs CIs that are more accurate and less biased than other bootstrapping methods since the skewness in distribution was handled poorly by all bootstrapping methods except bootstrap-*t* (Hoyle and Cameron 2003). For the confidence coefficient, the most typical 95% confidence interval ($\alpha = 0.05$) is applied. The coverage probability will be precise or close to $(1-\alpha)$, as is known if the data come from a symmetric distribution.

Based on Hoyle & Cameron (2003), bootstrap methods are thought to have two primary benefits over more conventional ways: (a) they are more robust, coping effectively with data that does not comply with standard parametric assumptions, and (b) they are frequently easier to enforce, relying on computer power to replace complex derivations. In additional, using bootstrap CI has the benefit of not requiring any assumptions about the distribution shape and universality of the approach (Klaudia and Łukasz 2020). Bootstrapping is a computer-intensive statistical technique that relies significantly on modern high-speed digital computers to do massive computations. Three phases can be utilized to generalize the bootstrapping method. To begin, resample using replacement samples to construct the bootstrap replications. Secondly, for every sample, compute the statistic of interest. Finally, make inferences based on the repeated statistic's distribution. To determine the sample mean, standard error, and generate CIs. The assumptions of the bootstrap method are the samples must be sufficient. If the sample sizes are too small, then the distributions cannot capture (Rousselet, Pernet, and Wilcox 2021). Additionally, for regression, hypothesis testing, adaptive estimation, calibration, and bioequivalence approaches, bootstrapping methodology is accessible. For the future work, do not use extreme values for bootstrap method (Onyesom and S.I. 2021).

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Acknowledgment

The authors would like to thank the Editor and anonymous reviewers for detailed reading that resulted in improved version of this paper.

References

[1] Abu-shawiesh, M. O. A., & Saeed, N. (2022). Estimating the population mean under contaminated normal distribution based on deciles.

[2] Abu-Shawiesh, M. O. A., Sinsomboonthong, J., & Kibria, B. M. G. (2022). A modified robust confidence interval for the population mean of distribution based on deciles. *Statistics in Transition New Series*, *23*(1), 109-128. https://doi.org/10.21307/stattrans-2022-007.

[3] Banjanovic, E. S., & Osborne, J. W. (2016). Confidence intervals for effect sizes: Applying bootstrap resampling. *Practical Assessment, Research and Evaluation*, *21*(5).

[4] Barker, N. (2005). A practical introduction to the bootstrap using the SAS system. *Proceedings of the Pharmaceutical Users Software Exchange Conference*, Paper PK02. http://www.lexjansen.com/phuse/2005/pk/pk02.pdf.

[5] Berrar, D. (2019). *Introduction to the non-parametric bootstrap*. April, 0-15. https://doi.org/10.1016/B978-0-12-809633-8

[6] Bickel, D. R., & Frühwirth, R. (2006). On a fast, robust estimator of the mode: Comparisons to other robust estimators with applications. *Computational Statistics and Data Analysis, 50*(12), 3500-3530. https://doi.org/10.1016/j.csda.2005.07.011.

[7] Boos, D. D., & Hughes-Oliver, J. M. (2000). How large does n have to be for z and t intervals? *American Statistician*, *54*(2), 121-128. https://doi.org/10.1080/00031305.2000.10474524.

[8] Chen, D., & Fritz, M. S. (2021). comparing alternative corrections for bias in the bias-corrected bootstrap test of mediation. evaluation and the health professions, *44*(4), 416-427. https://doi.org/10.1177/01632787211024356.

[9] Chou, C. Y., Lin, Y. C., Chang, C. L., & Chen, C. H. (2006). On the bootstrap confidence intervals of the process incapability index Cpp. *Reliability Engineering and System Safety*, *91*(4), 452-459. https://doi.org/10.1016/j.ress.2005.03.004.

[10] Cumming, G. (2007). Inference by eye: Pictures of confidence intervals and thinking about levels of confidence. *Teaching Statistics*, *29*(3), 89-93. https://doi.org/10.1111/j.1467-9639.2007.00267.x.

[11] Das, S. K. (2019). Confidence interval is more informative than P-Value in research. *International Journal of Engineering Applied Sciences and Technology, 04*(06), 278-282. https://doi.org/10.33564/ijeast.2019.v04i06.045

[12] David, H. A. (1998). Early sample measures of variability. *Statistical Science, 13*(4), 368-377. https://doi.org/10.1214/ss/1028905831.

[13] DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science*, *11*(3), 189-228.

[14] Desharnais, B., Camirand-Lemyre, F., Mireault, P., & Skinner, C. D. (2015). Determination of confidence intervals in non-normal data: Application of the bootstrap to cocaine concentration in femoral blood. *Journal of Analytical Toxicology, 39*(2), 113-117. https://doi.org/10.1093/jat/bku127.

[15] Doğan, C. D. (2017). Applying bootstrap resampling to compute confidence intervals for various statistics with R. Egitim Arastirmalari - *Eurasian Journal of Educational Research*, *2017*(68), 1-18. https://doi.org/10.14689/ejer.2017.68.1.

[16] Dogan, G. (2004). Confidence interval estimation in system dynamics models: bootstrapping vs. likelihood ratio method. *Proceedings of the 22nd International Conference of the System Dynamics Society*, 1-38.

[17] Efron, B. (1979). Bootstrap method_another look at Jackknife. *Breakthroughs in Statistics*, 569-593.

[18] Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, *68*(3), 589-599. https://doi.org/10.1093/biomet/68.3.589.

[19] Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*(397), 171-185.

[20] Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. https://doi.org/10.1201/9780429246593.

[21] Endo, T., Watanabe, T., & Yamamoto, A. (2015). Confidence interval estimation by bootstrap method for uncertainty quantification using random sampling method. *Journal of Nuclear Science and Technology, 52*(7-8), 993-999. https://doi.org/10.1080/00223131.2015.1034216.

[22] Flowers-Cano, R. S., Ortiz-Gómez, R., León-Jiménez, J. E., Rivera, R. L., & Cruz, L. A. P. (2018). Comparison of bootstrap confidence intervals using Monte Carlo simulations. *Water* (Switzerland), *10*(2), 1-21. https://doi.org/10.3390/w10020166.

[23] Good, P. (2005). *Multivariate analysis*. Springer.

[24] Hanna, J. P., Stone, P., & Niekum, S. (2017). Bootstrapping with models: confidence intervals for off-policy evaluation. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS, 1*(May), 538-546.

[25] Haukoos, J. S., & Lewis, R. J. (2005). Advanced statistics: Bootstrapping confidence intervals for statistics with “difficult” distributions. *Academic Emergency Medicine, 12*(4), 360-365. https://doi.org/10.1197/j.aem.2004.11.018.

[26] Hedges, S. B. (1992). The number of replications needed for accurate estimation of the bootstrap P value in phylogenetic studies. *Molecular Biology and Evolution, 9*(2), 366-369. https://doi.org/10.1093/oxfordjournals.molbev.a040725.

[27] Hongyi, Li, G. S., & Maddala. (2007). Bootstrapping time series. *Econometric Review*, *15*(2), 115-158.

[28] Horowitz, J. L. (2019). Bootstrap methods in econometrics. *Annual Review of Economics*, 11. https://doi.org/10.1146/annurev-economics-080218-025651.

[29] Hoyle, S. D., & Cameron, D. S. (2003). Confidence intervals on catch estimates from a recreational fishing survey: A comparison of bootstrap methods. *Fisheries Management and Ecology*, *10*(2), 97-108. https://doi.org/10.1046/j.1365-2400.2003.00321.x.

[30] Jung, K., Lee, J., Gupta, V., & Cho, G. (2019). Comparison of bootstrap confidence interval methods for GSCA using a Monte Carlo simulation. *Frontiers in Psychology, 10*(October), 24-25. https://doi.org/10.3389/fpsyg.2019.02215.

[31] Kennedy, K., & Schumacher, P. (1993). The introduction of bootstrapping techniques for finding confidence intervals using simulation. *Journal of Computers in Mathematics and Science Teaching*, 12(1), 83-92. http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ467667.

[32] Klaudia, J., & Łukasz, S. (2020). The use of the bootstrap method for the assessment of investment effectiveness and risk – the case of confidence intervals estimation for the Sharpe ratio and TailVaR. *Journal of Banking and Financial Economics, 1*/2020(13), 40-50. https://doi.org/10.7172/2353-6845.jbfe.2020.1.3.

[33] LaFontaine, D. (2021). The history of bootstrapping: tracing the development of resampling with replacement. *Mathematics Enthusiast*, 18(1-2), 78-99. https://doi.org/10.54870/1551-3440.1515.

[34] Mahmudah, U., Surono, S., Prasetyo, P. W., & Haryati, A. E. (2023). *Forecasting educated unemployed people in indonesia using the bootstrap technique*. *12*(July 2022), 171-182. https://doi.org/10.22103/jmmr.2022.19368.1239.

[35] Mesabbah, M., Rashwan, W., & Arisha, A. (2015). An empirical estimation of statistical inferences for system dynamics model parameters. Proceedings - Winter Simulation Conference, 2015-Janua, 686–697. https://doi.org/10.1109/WSC.2014.7019932

[36] Onyesom, C., & S.I., A. (2021). *Bootstrapping - An introduction and its applications in statistics*. *9*(3), 22-28.

[37] Pek, J., Wong, A. C. M., & Wong, O. C. Y. (2017). confidence intervals for the mean of non-normal distribution: transform or not to transform. *Open Journal of Statistics*, *07*(03), 405-421. https://doi.org/10.4236/ojs.2017.73029.

[38] Petty, M. D. (2012). Calculating and using confidence intervals for model validation. *Fall Simulation Interoperability Workshop 2012, 2012 Fall SIW*, 37-45.

[39] Puth, M. T., Neuhäuser, M., & Ruxton, G. D. (2015). On the variety of methods for calculating confidence intervals by bootstrapping. *Journal of Animal Ecology*, *84*(4), 892-897. https://doi.org/10.1111/1365-2656.12382.

[40] Rahmandad, H., Jalali, M. S., & Ghoddusi, H. (2013). Estimation of uknown parameters in system dynamics models using the method of simulated moments. 1212-1221. https://doi.org/10.1128/AAC.03728-14.

[41] Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2021). The percentile bootstrap: a primer with step-by-step instructions in R. *Advances in Methods and Practices in Psychological Science*, 4(1). https://doi.org/10.1177/2515245920911881.

[42] Saha, D., & Kapilesh, B. (2016). Estimation of concrete characteristic strength from limited data by bootstrap. *Journal of Asian Concrete Federation*, *2(*1), 81. https://doi.org/10.18702/acf.2016.06.2.1.81.

[43] Saito, Y., & Dohi, T. (2018). Parametric bootstrap methods for estimating model parameters of Non-homogeneous gamma process. *International Journal of Mathematical, Engineering and Management Sciences*, *3*(2), 167-176. https://doi.org/10.33889/ijmems.2018.3.2-013.

[44] Salkind, N. (2010). Encyclopedia of Research Design - Google Books. In SAGE Publications (p. 589). https://books.google.co.uk/books?hl=en&lr=&id=pvo1SauGirsC&oi=fnd&pg=PA1149&dq=Salkind,+N.+J.+(2010).+Encyclopedia+of+research+design+&ots=qte0QEw_k8&sig=r49o4D6rQ4VYpmpn_vmwMX9vGp8&redir_esc=y#v=onepage&q=inclusion&f=false%0Ahttps://books.google.com.ng/b.

[45] Shao, J., & Tu, D. (1995). *The jackknife and bootstrap*. Springer Science & Business Media. http://orton.catie.ac.cr/cgi-bin/wxis.exe/?IsisScript=BOSQUE.xis&method=post&formato=2&cantidad=1&expresion=mfn=004867%5Cnpapers2://publication/uuid/3565770D-2B58-427B-AFBC-BD0A21C7CC8D.

[46] Shi, W., & Golam Kibria, B. M. (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. https://doi.org/10.1080/00207390601116086.

[47] Sinsomboonthong, J., Abu-Shawiesh, M. O. A., & Kibria, B. M. G. (2020). Performance of robust confidence intervals for estimating population mean under both non-normality and in presence of outliers. *Advances in Science, Technology and Engineering Systems, 5*(3), 442-449. https://doi.org/10.25046/aj050355.

[48] Tapia, J. A., Salvador, B., & Rodríguez, J. M. (2020). Data envelopment analysis with estimated output data: Confidence intervals efficiency. *Measurement: Journal of the International Measurement Confederation*, *152*. https://doi.org/10.1016/j.measurement.2019.107364.

[49] Tong, L. I., Saminathan, R., & Chang, C. W. (2016). Uncertainty assessment of non-normal emission estimates using non-parametric bootstrap confidence intervals. *Journal of Environmental Informatics*, *28*(1), 61–70. https://doi.org/10.3808/jei.201500322.

[50] Tong, Lee Ing, Chang, C. W., Jin, S. E., & Saminathan, R. (2012). Quantifying uncertainty of emission estimates in National Greenhouse Gas Inventories using bootstrap confidence intervals. *Atmospheric*

Environment, *56*, 80-87. https://doi.org/10.1016/j.atmosenv.2012.03.063.

[51] Wang, F.-K. (2001). Non-normal data. *Quality and Reliability Engineering International*, *267*(259), 257-267. https://doi.org/10.1002/qre400.

[52] Wilcox, R. R. (2021). A note on computing a confidence interval for the mean. *Communications in Statistics: Simulation and Computation,* 1-3. https://doi.org/10.1080/03610918.2021.2011926.

[53] Yan, L. (2022). Confidence interval estimation of the common mean of several gamma populations. *PLoS ONE*, *17*(6 June), 1-13. https://doi.org/10.1371/journal.pone.0269971.

[54] Zaman, T. (2016). A study based on the application of bootstrap and jackknife methods in simple linear regression analysis. *International Journal of Sciences.*

[55] Zhou, X. H., & Dinh, P. (2005). Nonparametric confidence intervals for the one- And two-sample problems. *Biostatistic*s, *6*(2), 187-200. https://doi.org/10.1093/biostatistics/kxi002.

[56] Zhao, S., Yang, Z., Musa, S. S., Ran, J., Chong, M. K. C., Javanbakht, M., He, D., & Wang, M. H. (2021). Attach importance of the bootstrap t-test against Student's t-test in clinical epidemiology: A demonstrative comparison using COVID-19 as an example. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268821001047x.