Bootstrap-Based Confidence Intervals: Methods, Applications, and Challenges

Yahan Zhang *

School of Arts and Sciences, Rutgers University, New Brunswick, the United State

* Corresponding Author Email: yz1932@scarletmail.rutgers.edu

Abstract. Bootstrap methods offer a flexible and generally applicable framework for statistical inference, especially when classical parametric assumptions such as normality and large-sample approximations are violated. This paper offers a systematic review of confidence intervals constructed via bootstrap methods, including their conceptual foundations, methodological variants, and applications. After presenting an overview of the motivation for bootstrap confidence intervals and their advantages over classical analytical methods, three major methods are reviewed and discussed: the percentile method, the bias-corrected and accelerated method, and the bootstrap-t method. These methods are compared and evaluated according to their bias correction, complexity, and accuracy. Some applications of these methods, such as group comparisons, correlation coefficient estimation, and mediation effects, are reviewed and discussed to illustrate their applicability and effectiveness under non-normal and complex data conditions. Finally, some methodological challenges and limitations, such as small sample size, method selection, and computational issues, are reviewed and discussed. By synthesizing recent research and applied studies on confidence intervals via bootstrap methods, this review article attempts to highlight the applicability and relevance of confidence intervals via bootstrap methods as a robust and flexible statistical tool for modern statistical practices.

Keywords: Bootstrap; Confidence Interval; Resampling; Statistical Inference.

1. Introduction

The process of statistical inference is said to be a part of any kind of research that is considered to be empirical in nature. This is because the process of statistical inference is considered to be quantification of the uncertainty associated with the parameter of interest. In general, in any kind of research that is considered to be empirical in nature, it is said that the aim of the researcher is not only to obtain a certain number of values but also to understand the reliability associated with the values obtained. The confidence intervals are said to be a part of this process, and this is considered to be more informative compared to single-point estimates.

Usually, the calculation of CIs is done via parametric methods. For example, in classical methods, normality or large sample sizes are assumed. However, in real-world problems, such conditions are not met. For example, in real-world problems, the data distribution is not symmetric and has heavy tails and/or outliers. Furthermore, the sample sizes in real-world problems are small due to certain constraints (Chernozhukov et al., 2023) [2]. Thus, the validity of traditional parametric methods in real-world problems for calculating CIs is questionable. This can lead to incorrect results in statistics.

In order to overcome the shortcomings described above, the resampling-based methods have been extensively studied and developed in statistical inference. Among these resampling-based methods, the bootstrap method developed formally by Efron and Tibshirani is a powerful tool in statistical inference (Tibshirani and Efron, 1993) [9]. The bootstrap method is based on the idea that the original data set is used to resample the sampling distribution of a statistic. By using the original data distribution instead of the real population distribution, the bootstrap method could potentially be used to facilitate statistical inference.

Bootstrap confidence intervals are considered to be one of the most important uses of the bootstrap technique (Cruz et al., 2024) [4]. Unlike the conventional method of obtaining confidence intervals, which is based on a theoretical distribution, the confidence intervals are obtained using the empirical distribution. This helps the researcher to draw inference for a wide variety of estimators for which

drawing confidence intervals is not feasible or is difficult. In particular, the confidence intervals have been observed to be highly effective in handling non-normal data, small samples, etc.

The applicability of bootstrap CIs extends far beyond the simplicity of the method itself. In fact, data analysis in the last few years has been based on sophisticated data structures and computationally intensive models, for which traditional methods of inference have been less reliable. The bootstrap methods have been shown to offer a useful framework for uncertainty analysis, which has gained popularity owing to its robustness to distributional misspecification and its ability to cover a wide range of statistical quantities of interest. Bootstrap CIs have been used in different areas such as psychology, biology, economics, and social sciences.

Despite the fact that bootstrap CIs are very versatile, they may not always be the best. For example, their accuracy may not be good when the samples are very small. Besides that, different bootstrap CIs may produce very different results. This again points to the importance of understanding the theory behind bootstrap CIs rather than simply using them.

On the basis of the aforementioned background, the major purpose of this paper is to propose a structured review on the methods of bootstrap confidence intervals in terms of their conceptual foundations, methodological developments, and applications. Rather than focusing on the mathematical derivations of the methods of bootstrap confidence intervals, the major purpose of the proposed review is to elucidate the conceptual foundations of the most popular methods of bootstrap confidence intervals and assess the advantages and disadvantages of the methods of bootstrap confidence intervals in terms of their applications. The structure of the proposed paper is as follows: In the following section, the latest research on the methods of bootstrap confidence intervals will be reviewed. In the third section, the major methods of bootstrap confidence intervals will be introduced, and the characteristics of the methods of bootstrap confidence intervals will be discussed. In the fourth section, the representative applications of the methods of bootstrap confidence intervals will be presented and evaluated. In the fifth section, the current challenges and future research on the methods of bootstrap confidence intervals will be discussed. Finally, in the sixth section, the summary and outlook of the proposed paper will be provided.

2. Related work

In their tutorial paper, Rousselet et al. present an introduction to the percentile bootstrap method with particular emphasis on providing descriptions that are easily understandable and implementing the bootstrap method step-by-step. Rousselet et al.'s paper illustrates the construction of bootstrap CI without assuming a specific distribution and using it for non-normal distributions and robust estimators (Rousselet et al., 2021) [8].

A comprehensive review of various bootstrap techniques for CI development is provided in the study done by Mokhtar et al. (Mokhtar et al., 2023) [7], which includes percentile methods, bias-corrected and accelerated (BCa), and bootstrap-t methods.

Practical implementation issues are also extensively discussed by Cheung et al., who provide a description of the application of constructing non-parametric bootstrap CIs for a broad range of statistics using standard statistical software (Cheung et al., 2023) [3].

This can be demonstrated with the use of applications in mediation analysis, which showed that the use of bootstrap CIs results in better performance compared to the use of traditional normal theory-based CIs, especially when the data are not normally distributed (Alfons et al., 2021) [1]. In addition, the use of the method for improving inference for non-normal data with unequal variances was demonstrated.

3. Main Bootstrap Confidence Interval Techniques

There are several methods for the construction of Bootstrap CIs, each based on the empirical distribution resulting from resampling with replacement. Although the techniques vary with respect

to the calculation of endpoints of the CI and the level of sensitivity to bias, skewness, and sample size, they employ a common technique of resampling. This section covers the percentile method, the BCa method, and the bootstrap-t method.

3.1. Percentile Bootstrap Confidence Interval

One of the simplest and most intuitive bootstrap CI methods is the percentile bootstrap CI. The procedure for this CI is: resample the given data with replacement and then calculate the statistic of interest for each resampled data set. This is the bootstrap distribution for the statistic. Then the confidence interval is computed by finding the appropriate quantile from the bootstrap distribution, such as the 2.5th and 97.5th percentiles for a 95% CI.

One of the major advantages that the percentile method enjoys is that it is conceptually clear. Unlike the other methods that rely on the theoretical results about the sampling distribution, the percentile method uses the actual bootstrap distribution to arrive at the confidence limits. This method has been pointed out by Rousselet et al. to be useful in obtaining confidence intervals for different statistics, including those for which the distribution is not known. Furthermore, the percentile bootstrap method is very simple to apply and assumes very few things besides the actual sample.

Nevertheless, the percentile method of constructing the CI has some disadvantages, some of which include the fact that the percentile method of constructing the CI can have an equal level of bias as the original estimator, given that the method is based on the empirical distribution of the bootstrap estimates. In particular, the percentile method of constructing the CI can suffer from poor coverage accuracy when the distribution of the bootstrap estimates is skewed or shifted compared to the true distribution. In fact, the percentile method of constructing the CI can have a level of bias, despite the fact that the method performs fairly well, as suggested by Mokhtar et al. (Mokhtar et al., 2023) [7].

Despite the aforementioned disadvantages, the percentile bootstrap method remains popular owing to its simplicity and wide applicability. This method is usually advocated as a first option for bootstrap-based inference.

3.2. Bias-Corrected and Accelerated (BCa) Bootstrap Confidence Interval

The bias-corrected and accelerated bootstrap confidence interval method is an extension of the percentile method. This method attempts to correct for bias and skewness in the bootstrap distribution. In addition to the resampling method in the percentile method, this method uses two additional parameters: the bias correction and acceleration. These parameters improve the estimates of the actual distribution.

The rationale for the use of the intervals can be derived from the fact that the coverage accuracy can be improved, especially when the model is biased. In such situations, the use of the intervals can yield effective results compared to the use of percentile intervals. In the research paper on the topic, Mokhtar et al. note that the use of confidence intervals can yield effective results compared to the use of percentile intervals.

While it is more accurate, the BCa method also introduces more complexity in the computation. In fact, the computation for the acceleration parameter often involves computations using jackknife resampling techniques. Moreover, the interpretation of BCa intervals is not as intuitive compared to percentile intervals.

In actual fact, the BCa bootstrap CI is a compromise between accuracy and complexity. Although the BCa bootstrap CI might be more effective compared to the percentile interval, it also needs more careful implementation. Therefore, the BCa bootstrap CI is appropriate for situations where the need for accurate inference is a necessity, and computation is not a problem. Although the coverage probabilities of the BCa bootstrap CI are better compared to the percentile interval, the computation of the BCa bootstrap CI is more complicated because of bias and skewness.

3.3. Bootstrap-t Confidence Interval

The confidence interval for the bootstrap-t method takes a different approach in that the results are standardized using an estimate of the standard error. Rather than resampling the statistic, the method uses a distribution of studentized statistics to arrive at the confidence limits. In this process, the confidence interval is obtained.

The main advantage that the bootstrap-t method has over the other methods from a theoretical point of view is that the variation of the results can be controlled through standardization, which could result in a better outcome. In ideal situations, the bootstrap-t method could result in a better coverage of the intervals compared to the percentile method. However, the main disadvantage that the bootstrap-t method has is that the results depend on the standard error. In case the standard error has not been well estimated, the results may not yield the desired outcome.

Practically, bootstrap-t methods can be more computationally intensive to obtain than percentile or BCa intervals. This is due to the additional resampling and estimation of variance that is needed to obtain a bootstrap-t interval, which can be problematic for complex statistics and small sample sizes. As such, although bootstrap-t methods are theoretically attractive, they are not of great applied interest.

Cheung et al. stress that the selection of bootstrap CI methods should be based on the features of the data, the complexity of the estimator, and the limitations (Cheung et al., 2023) [3]. Even though bootstrap-t methods have the potential to attain better accuracy than percentile and BCa methods, the latter have a higher chance of being used in applied research due to their simplicity. Bootstrap-t methods have better accuracy than percentile and BCa methods in theory but are not likely to be used in practice due to their complexity and the difficulties in obtaining a stable estimate of variance.

4. Applications and Result Analysis

Bootstrap confidence intervals (CIs) have proven useful tools for empirical research, especially when faced with situations where the assumptions of classical parametric methods are not met, or when the estimators have complicated sampling properties. In the following, the paper shall discuss some representative uses of bootstrap CIs, together with the benefits of such methods through the use of case studies.

One of the relevant application domains is the one associated with the inference under non-normal data and/or non-homogeneous variances. It is shown by Johnston and Faulkner that, with regard to the comparison of groups with non-normal data and/or non-homogeneous variances, the performance of the methods based on the bootstrap is superior compared to the other methods (Johnston and Faulkner, 2021) [6]. In this particular domain, the level of reliability associated with the uncertainty evaluation of the CIs developed using the bootstrap is higher compared to the traditional ones.

Bootstrap CIs have been found to be effective in dealing with uncertainty related to correlation measures. DelosReyes and Padilla consider the issue of bootstrap confidence intervals for correlation coefficients (DelosReyes and Padilla, 2024) [5]. They assess the coverage accuracy of such confidence intervals under various distributional assumptions. It appears from their study that better accuracy can be achieved for bootstrap confidence intervals, especially when the distribution is symmetric. This study points to the versatility of bootstrap CIs in dealing with dependence measures and their potential to handle statistical quantities whose sampling distributions are difficult to approximate.

For more complicated inferential problems, the use of bootstrap CIs provides a promising avenue for the assessment of uncertainty for non-standard estimators. Alfons et al. provides a recent example of the application of bootstrap methods for the assessment of mediation effects, where the distribution of indirect effects is usually asymmetric and non-normal (Alfons et al., 2021) [1]. Their work demonstrates the superiority of bootstrap CIs over the standard approach, especially when dealing with the effects of outliers and non-normality of the data distribution.

However, it appears that a common factor is present among each of these problems: where the assumptions were violated, the confidence intervals using the bootstrap method were found to be more robust. The versatility of the bootstrap method in handling problems with irregularities such as non-normality and non-constant variance is one of the strengths of the bootstrap method. Overall, the above problems presented a discussion on the versatility of the confidence intervals using the bootstrap method. Bootstrap confidence intervals were effective in handling problems such as comparing groups and correlation, as well as the effect of mediation.

5. Discussion

5.1. Challenges

While they are highly flexible and applicable, bootstrap confidence intervals are not without their limitations. One of the major limitations of these confidence intervals is related to small sample sizes. This is because, due to their empirical distribution, small sample sizes are often considered to be a poor estimate of the population.

Another challenge associated with bootstrap CIs is the method selection. As indicated in the previous sections, various bootstrap methods for constructing CIs, such as percentile, bias-corrected and accelerated (BCa), and bootstrap-t, may produce very dissimilar results depending on the nature of the data. As indicated by Mokhtar et al., no single bootstrap method for constructing CIs is better in all situations (Mokhtar et al., 2023) [7].

Computational cost is another concern, even if the current resources available have made resampling methods more accessible, since the bootstrap methods require a considerable number of resamples for the computation of the results. In addition, for the application of the BCa and bootstrap-t methods, a resampling process for the computation of the variance needs to be considered, which adds more burden for the computation of the results.

Finally, it is noted in the current discussions that bootstrap CIs should not be thought of as a general solution to addressing uncertainty quantification. In particular, Zrimšek and Štrumbelj note that, while bootstrap methods represent a powerful general framework, their effectiveness ultimately depends on the extent to which the empirical distribution is assumed to represent the true data-generating process, with a violation of this assumption leading to potentially misleading results (Zrimšek and Štrumbelj, 2025) [10].

5.2. Future Directions

The focus of the bootstrap confidence intervals in the near future is expected to be on the extension of the bootstrap confidence intervals to more complex data environments. An area that is expected to be focused on more in the near future is the extension of the bootstrap techniques to the high-dimensional data environments and the machine learning models, where the traditional statistical methods do not produce satisfactory results. The problem that arises with the extension of the bootstrap techniques is the identification of the computationally efficient bootstrap techniques.

Another area of interest in this case relates to the improvement of the accuracy of coverage with the use of hybrid or adaptive techniques. For instance, the use of the bootstrap techniques along with robust estimates and other resampling techniques may potentially improve the results, reducing the effects of outliers and irregularities. Another area of interest relates to the further exploration of the method selection strategies for the selection of the appropriate bootstrap CI techniques.

From the applied perspective, more emphasis on the provision of guidelines might help to improve the application of the results obtained using the bootstrap CIs. For instance, providing some guidelines on how to determine the sample sizes might help to bridge the gap between the development of the theory and the results obtained.

In summary, future methodological development, as well as progress in computer power, is expected to consolidate the position of bootstrap confidence intervals in contemporary data analysis.

Bootstrap methods are expected to overcome the current limitations and difficulties to remain a key technique in the face of increasingly complex research conditions.

6. ConclusionS

Bootstrap confidence intervals have been established as a vital technique in carrying out statistical inferences, especially in situations where classical parametric methods are either not reliable or not appropriate. The techniques that have been developed from the application of the bootstrap method have been established to be a promising approach in making statistical inferences. This paper discussed the major techniques that are used in making confidence interval constructions using the bootstrap method. Furthermore, this paper sought to explore the concept, advantages, and disadvantages associated with the techniques.

From the analysis, it can be concluded that different techniques have different strengths. In the percentile method, the simplicity of the method makes the percentile method the best for exploratory analysis. In the BCa method, the accuracy of the results can be improved, though the process is computationally intensive. In the bootstrap-t method, standardization can be integrated, and the method can produce strong theoretical results, though the process can be complex. From the results, the effectiveness of the bootstrap CIs can be seen for different conditions, such as when comparing groups with non-normal distributions, correlation analysis, and when the sampling distribution is asymmetrical.

Of course, while doing all of this, certain limitations have to be kept in mind. It is a known fact that bootstrap methods are unstable for small samples. In addition, the results of various methods of constructing confidence intervals may differ. This proves that bootstrap methods have to be used with caution.

If people look to the future, it seems safe to say that additional refinements in bootstrap techniques and the development of guidelines will continue to heighten the importance of results obtained via the bootstrap method. In fact, with additional computer power and sophisticated data structures, bootstrap confidence intervals may become an even more critical aspect of statistical research in the future. Bootstrap confidence intervals continue to represent a powerful and flexible tool in statistical research

References

[1] Alfons, A., Ateş, N. Y., & Groenen, P. J. 2021. A robust bootstrap test for mediation analysis. Organizational research methods, 25(3), 591-617.

[2] Chernozhukov, V., Chetverikov, D., Kato, K., & Koike, Y. 2023. High-dimensional data bootstrap. Annual Review of Statistics and Its Application, 10(1), 427-449.

[3] Cheung, S. F., Pesigan, I. J. A., & Vong, W. N. 2023. DIY bootstrapping: Getting the nonparametric bootstrap confidence interval in SPSS for any statistics or function of statistics (when this bootstrapping is appropriate). Behavior Research Methods, 55(2), 474-490.

[4] Cruz, L., Blanco, J., & Giraldo, R. 2024. Bootstrap versus Jackknife: Confidence intervals, hypothesis testing, density estimation, and kernel regression. Ciencia en Desarrollo, 15(2), 143-152.

[5] DelosReyes, J. M. V., & Padilla, M. A. 2024. Bootstrap correlation confidence interval estimation: the positive impact of a symmetric distribution. The Journal of Experimental Education, 92(3), 559-579.

[6] Johnston, M. G., & Faulkner, C. 2021. A bootstrap approach is a superior statistical method for the comparison of non-normal data with differing variances. The New Phytologist, 230(1), 23-26.

[7] Mokhtar, S. F., Yusof, Z. M., & Sapiri, H. 2023. Confidence intervals by bootstrapping approach: a significance review. Malaysian Journal of Fundamental and Applied Sciences, 19(1), 30-42.

[8] 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), 2515245920911881.

[9] Tibshirani, R. J., & Efron, B. 1993. An introduction to the bootstrap. Monographs on statistics and applied probability, 57(1), 1-436.

[10] Zrimšek, U., & Štrumbelj, E. 2025. Quantifying uncertainty: All we need is the bootstrap? Journal of Statistical Computation and Simulation, 1-19.