EMPIRICAL INVESTIGATION ON THE EFFECT OF THE NUMBER OF RESAMPLINGS ON THE DISTRIBUTION OF BOOTSTRAP STANDARD ERROR USING RESPONSE TIME DATA

ABSTRACT.

Aims: Toinvestigate how the number of bootstrapping B affects the values returned by the bootstrap standard error of the arithmetic mean and the α-trimmed mean of response data using bootstrap confidence intervals (CI) at 95% level; carried out to fill up observed gap for study on standard error, the tool generally employed in assessing the long run accuracy of a given statistical estimator of θ.

Study design: **This was a parametric,** empirical bootstrap simulation study.

Place and Duration of the study: Departments of Computer Science and Statistics, Federal Polytechnic Oko, 2020/2021 session.

Methodology: Response time data were generated with studentcustomers of mobile telephone network (mtn) Nigeria and stored in SPSS.A sample n = 51 responses wasselected using "Select Cases" command to increase precision and minimize bias. Bootstrap simulation study was carried out using R programming language. Four approaches for estimating bootstrap confidence intervals were used. Theinterval coverage and the interval lengths were determined and compared for B = 20, 50, 100, 500, 1000, 5000, and 10000.

Results: The 95% CI for s_{10} = 0.2266338 (the estimated sample standard error of 10% trimmed mean) returned the best interval for our skewed data set; when B = 20; the CIfor s_{10}returned (0.2051, 0.2343) for the normal approach, (0.2082, 0.2371) for the basic, (0.2071, 0.2360) for the percentile and (0.2071, 0.2360) for the BCa method. As B increased to 5000, itreturned (0.2259, 0.2277) for the normal approach, (0.2259, 0.2277) for the basic, (0.2260, 0.2277) for the percentile and (0.2261, 0.2279) for the BCa showing a shorter interval yet covering the estimate.

Conclusion: Thusfor our response data study, increasing B in estimating standard error increases the chances of more precise and shorter confidence intervals rather than the chances for coverage.

Key words: *Bootstrap, standard error, trimmed mean, number of bootstrapping, response data.*

1 INTRODUCTION:

A bootstrapmethod is a resampling technique often employed in the construction of confidence intervalsto estimate the variation of point estimates. The approach is largely based on the law of large numbers (Orloff, and Bloom,2017) [1].**Dixon** (2006) [2] noted that at the heart of the bootstrap method is the concept that thedistribution of the statistic of interest can beapproximated by estimates from repeated samples.The bootstrap technic makes use of the information available in the data without making any adventitious assumptions about the distribution of the population of the data. The bootstrap is a general methodology for answering question on the accuracy of an estimator of θ and the standard error is often used to measure long run accuracy of such estimator.A concern however is often on the number of resampling that will yield the optimum estimate of interest especially for asymmetric data set.Most available works have been about the mean and other measures of the center than about the standard error. (Efron and Tibshirani(1986)[3], Davison and Hinkley (1997) [4],Davidson and Mackinnon (2000)[5],Wei,(2013) [6],Rousselet, Pernet, and Wilcox, (2019)[7],Diez,(n.d.),[8],Hesterberg (2014)[9], Robert (2019)[10])

The distributions of response time (RT) data violates the normality assumption underlying classical statistical analyses because they are generally positively skewed and unimodal in shape (Heathcote, Popiel, and Mewhort, 1991)[11]. Customers of service providers in Nigeria of which the researcher is one, often experience delays from the service providers in responding to their online complaints via their online customer care. This can be frustrating most times. Service providers' response to online customers'complaint is a response time data set and as such should be expected to be skewed.

According to Chernick and LaBudde, (2011)[12], the arithmetic sample mean is the theoretical mean of the bootstrap distribution. Lampert, Stahel and Abbt (2001)[13] opined that response time data distributions are mostly skewed and the arithmetic mean generally delivers a poor measure of central tendency for skewed distributions. The α-trimmed mean provides more accurate information about the central tendency than the mean for an asymmetric distribution. Since the sample mean is the theoretical mean of the bootstrap distribution and the α-trimmed mean is the preferred mean for skewed data, we decided to study the bootstrap standard error of the mean as well as that of the α-trimmed mean. Writing on the bootstrap confidence interval coverage,Rousselet, Pernet, and Wilcox, (2021)[14]and Rousselet, and Wilcox, (2019)[15]noted that when sampling from skewed data, a 95% bootstrap percentile confidence interval for the mean can actually yield a 88% interval while that of 20% trimmed mean may yield as much as 94.6% coverage.

Our interest was to investigate how the number of resampling affects the values returned by the bootstrap standard error of the arithmetic mean and that of the α-trimmed mean using bootstrap confidence interval at 95% level.

Writing on the needed number of bootstrap samples, Davidson and Mackinnon (2000)[5], Orloff and Bloom (2017)[1] believe that B the number of bootstrap resampling can influence the accuracy of the bootstrap estimate. According to Davison, Hinkley, and Young (2003) [16] typically to achieve a negligible simulation variation, a few thousand bootstrap samples are needed.Specifically in the opinion of Davidson and Mackinnon (2000)[5], where the cost of computation is affordable, B should be extremely large.On the basis of the above opinions, researchers may spend much resources trying to run large number of resampling in setting up bootstrap confidence intervals hoping possibly to increase the chances of the interval including the estimate of interest.

Four different approaches are considered to estimating bootstrap confidence interval namely: the first order normal approximation, the basic bootstrap interval, the bootstrap percentile interval, and the adjusted bootstrap percentile that is the bias-corrected accelerated(BCa) interval. These were used to set up the confidence intervals for the selected measures. How these different approaches behave in the estimation of the confidence interval of the standard error for the mean and trimmed mean for skewed response time data and as the number of bootstraps increased was considered.

Efron and Tibshirani(1986)[3] opined that for estimating the standard errors and coefficient of Variations (CV), B as small as 25 gives reasonable results. They iterated however that the situation is different for the construction of confidence intervals. Efron and Tibshirani[3] quoting the calculations of Efron (1984)[17] stated that B =1000 is a rough minimum for the number of Monte Carlo bootstraps necessary to compute BC or BCa. They maintained however that somewhat smaller values like B = 250 can give a useful percentile interval. Davison and Hinkley (1997)[4] recommended the number of bootstraps B = 999 for confidence interval levels of 95% and 99% where such is practically feasible. Using computer packages today, it will take afew seconds to reach such number. The question however is how relevant is it to keep B increasing irrespective of the shape of the data set?

From the fore going therefore, it is expected that the more the number of resampling B, the more accurate the estimate and the shorter the confidence interval.It is therefore expected that the interval will be getting shorter as B gets larger and with anincreased probability of coverage.Using the service provider data therefore and employing the confidence interval tool, we studied how B affects the accuracy and the precision of the estimates of bootstrap standard error for the selected bootstrap methods for calculating CI at 95% level.

The bootstrap standard error of the arithmetic mean as well as that of the α-trimmed mean and their 95% confidence intervals were estimated using the four approaches and compared. The interval lengths were equally calculated and compared. The study covered for B (the number of bootstrap resampling) = 20, 50, 100, 500, 1000, 5000, 10000.

2. MATERIALS AND METHODS.

2.1 Data Collection Procedure

The data for this research was obtained fromStatistics and Computer Science students of Federal Polytechnic Oko Atani campus, Anambra state Nigeria. The selected students were all customers of Mobile Telephone Network (MTN) Nigeria; oneof the major service providers in the southeast of Nigeria. Volunteer participants were utilized to obtain the data. In the course of obtaining the data, the participants were asked to place calls to their network provider MTN using thenetwork's call center number at separate times. The intention of a caller/customer in using online customer care is always to speak with the online customer care representative within the shortest possible time. Oftentimes however, when calls are placed on them, some messages are being played in the form of adverts and jingles before transferring the caller to a customer care representative. Some network providers will display their adverts and jingles for a long time before finally allowing the caller/customer to speak with an agent. This could be very frustrating as well as time and money- consuming. I personally do not like that and it often discourages customers from utilizing such services.

The participants were told to call their customer care line and follow the required processes until they are transferred to a customer care agent to attend to them. They were to end the call immediately afterthe customer care representative responds to the call and record the total time spent in minutes. These were done severally for different periods and the time/duration of the calls recorded. In the end, the data wereretrieved, organised and entered in the Statistical Product and Services Solutions (Hejase and Hejase, 2013)[18] IBMSPSS and the desired sample was randomly selected using the “Select Cases” command from the data menu with the purpose to increase precision and avoid bias.

2.2Theory and Methods.

2.2.1 The Bootstrap Principle and Procedure.

The bootstrap principle indicatesthat the empirical distribution obtained by resampling from a sample is approximately equal to the theoretical or true distribution (Orloff, and Bloom; 2017)[1].

Letθ = any statistics computed from the original sample data;

θ^* = a statistic by the same formula as θ but computed instead using a resampled data from the original sample.

To calculatethe desired bootstrap estimate θ^*, we apply the following procedure:

(1) Select a sample of size n;$x_1, x_2, x_3, \dots x_n$from a distribution F to form an empirical distribution $F_n$which puts equal mass $\frac{1}{n}$ on the selected n data points.

(2) Calculate the desired estimate θ.

(3) Select a sample of size $n = x_1^*, x_2^*, x_3^*, \dots x_n^*$with replacement from the original sample in(1) above

(4) Calculate θ_i^* the value of θ from the ith resampled data using the same formula for calculating θ.

(5) Repeat (3) and (4) B times so that we have $\theta_1^*, \theta_2^*, \theta_3^*, \dots, \theta_B^*$

(6) The bootstrap estimate of θ, θ_b is obtained by calculating θ still using the same formula but based on$\theta_1^*, \theta_2^*, \theta_3^*, \dots, \theta_B^*$. (Wilcox, 2005)[19]

2.2.2 Bootstrap Estimate of the Standard Error of the Mean.

The long run accuracy of a given statistical estimator of θ is usually assed with its standard error(Staudte and Sheather,1990)[20]

The standard error of the arithmetic mean given n samples is given as

$s_{\bar{x}} = \sqrt{\frac{\sum_i^n (x_i - \bar{x})^2}{n-1}}$ (Madsen and Moeschberger,1986)[21]

Following the procedure for bootstrap estimates presented above, the bootstrap estimate for the standard error of the standard error of arithmetic mean is given as: $s_b = \sqrt{\frac{\sum_{i=1}^{B}(s_i^* - \bar{s}^*)^2}{B-1}}$: (1)

$Where\ s_b = the\ bootstrap\ standard\ error\ of\ the\ standard\ error.$

$s_i^* = the\ standard\ error\ of\ the\ ith\ bootstrap\ sample,$

$\bar{s}^* = \frac{\sum_{i=1}^{B} s_i^*}{B}$ $=the\ arithmatic\ mean\ of\ the\ standard\ error\ of\ the\ B\ bootstrap\ sample$

$standard\ errors$ andB= the number of bootstrap samples.

2.2.3 Bootstrap Estimate of the Standard Error of the α-Trimmed Mean.

The standard error of the α-trimmed mean is given as:$s_\alpha = \sqrt{\frac{s_w{}^2}{(1-\alpha)^2 n}} = \frac{s_w}{(1-\alpha)\sqrt{n}}$ (2)

$Where s_w{}^2 = \frac{1}{n-\alpha}\sum_{i=1}^{n}(w_i - \bar{w})^2 = sample\ winsorized\ variance$(Staudte and Sheather, 1990,[20] Wilcox; 2005 [19])

The bootstrap version$s_{b\alpha} = \sqrt{\frac{s_{bw}^2}{(1-\alpha)^2 B}} = \frac{s_{bw}}{(1-\alpha)\sqrt{B}}; where\ s_{bw}^2 = \frac{1}{B-\alpha}\sum_{i=1}^{B}(w_i - \bar{w})^2$(3)

The limit cases of $\alpha = 0\ and\ \alpha = 0.5$ corresponds to the sample mean and median respectively (Staudte and Sheather, 1990[12], Maronna*et al*, 2006[14]). The choice of α is of practical importance. It is generally chosen so that $\bar{x}_\alpha$ is apt to have a relatively small standard error $\sqrt{var(\bar{x}_\alpha)}$ among commonly occurring situations.

Wilcox (2005) [19] noted that if α is too small, the standard error of the trimmed mean can be drastically inflated by outliers or by sampling from a heavy - tailed distribution. When one is sampling from a normal distribution ifthepercentage of trimming α is quite large, the standard error of the trimmed mean can be relatively large when compared to that of the arithmetic sample mean

2.3Method of Comparison. The method of comparison employed for the study is confidence interval(CI). CI is used to measure the precision of an estimate, to know the plausible range of values given the parameter of interest. It is the probability that the computed interval will include the population parameter (Efron & Tibshirani, 1993)[22].There are several available approaches for estimating bootstrap confidence intervals but as earlier stated, for this study four methods were considered. These are the first-order normal approximation, the basic bootstrap interval, the bootstrap percentile interval, and the adjusted bootstrap percentilewhich is the bias-corrected accelerated (BCa) interval.(Canty,2021) [23]

The first- order normal approximation is calculated using the normal approximation. The basic bootstrap interval uses the estimated standard error and is calculated using the basic bootstrap method, the bootstrap percentile interval and the BCa intervals use thebootstrap percentile method but the BCa interval in addition is also adjusted to account for bias and skewness (Larget, 2014) [24].

A95% confidence interval was set up using the four approaches named above.

We compared the confidence interval coverage of the selected procedures for B, (*the number of bootstrap resampling)* = 20, 50, 100, 500, 1000, 5000, and 10000.

The length of each confidence interval was determined as:

$$length = \hat{\theta}_{bup} - \hat{\theta}_{blo} \ \ where\ \hat{\theta}_{bup} = upper\ value\ of\ the\ bootstap\ interval \hat{\theta}_{blo} = the\ lower\ value\ of\ the\ bootstrap\ interval$$

Dixon(2006)[2] opined that apercentile interval may not have thecorrect coverage for a skewed sampling distribution.Efron (1984)[17] showed that B=1000 is a rough minimum for the number of Monte Carlobootstraps needed to compute BCa intervals. Efron (1987)[25] stated that the BCa approach to calculating the bootstrap confidence interval properties that put it at advantage over the other methods.The coverage errors for the confidence interval for the mean and the median is usually small using the BCa method whether the population distribution is normal or non- normal even for asample size of about 20 or more. (Efron andTibshirani, 1993[22]; Lei and Smith, 2003[26]).

Setting up a (1-α)% confidence interval implies that on average one would expect (1-α)% of the intervals to include the true value of the estimate while α% would not include it (Glasserman, 2001[27], Helwig, 2017[28]).

The R programming language was employed both for the simulation and analysis.

3. RESULTS AND DISCUSSIONS.

The data for this research work is presented on a histogram below. As should be expected from our literature review, the data is positively skewed beingresponse data(figure 1).

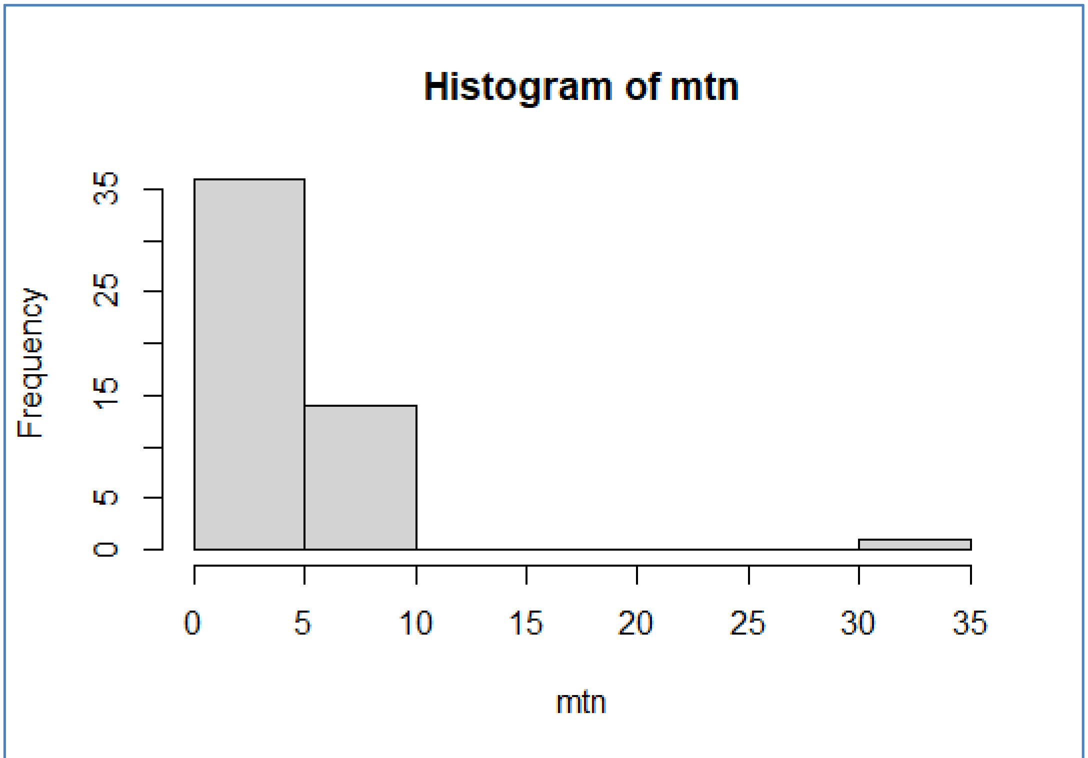

Fig 1. Histogram of the Response Data Used for the Study

From the analysis, the estimated sample standard error $s_{\bar{x}}$of the arithmetic mean for the n=51 set of data studied was4.123277,that of 10% trimmed mean $s_{10}$was 0.2266338: and s_{20} = 0.2663485 for the sample standard error of 20% trimmed mean.Thus: the α - trimmed mean returned smaller standard error than the arithmetic mean with the 10% trimming returning slightly smaller value than 20% trimming this could be possibly explained by the non-symmetric nature of our data set.

Table 1 is a summary of the resultsof the analysis while Table 2 is a display of the width of the bootstrap confidence intervals.

Table 1: 95% Confidence Intervals for the Bootstrap standard Errors($s_{\bar{x}}$, s_{10} and s_{20})

	Standard Error of the Mean($s_{\bar{x}}$)				Standard Error of the 10% Trimmed Mean(s_{10})				Standard Error of the 20% Trimmed Mean (s_{20})			
B	FONA	BBI	BPI	BCa	FONA	BBI	BPI	BCa	FONA	BBI	BPI	BCa
20	2.762, 4.656	2.651, 4.451	2.933, 4.733	2.933, 4.512	0.2051, 0.2343	0.2082, 0.2371	0.2071, 0.2360	0.2071, 0.2360	0.2246, 0.2507	0.2262, 0.2482	0.2263, 0.2482	0.2263, 0.2482
50	2.988, 4.355	2.890, 4.446	2.909, 4.465	2.939, 4.507	0.2043, 0.3204	0.1820, 0.2961	0.2277, 0.3418	0.2324, 0.3503	0.2514, 0.2649	0.2486, 0.2643	0.2514, 0.2671	0.2525, 0.2686
100	3.182, 3.910	3.201, 3.909	3.113, 3.822	3.124, 3.830	0.2201, 0.2293	0.2196, 0.2292	0.2206, 0.2302	0.2197, 0.2288	0.2412, 0.2603	0.2416, 0.2603	0.2409, 0.2595	0.2400, 0.2589
500	3.536, 3.852	3.529, 3.850	3.543, 3.863	3.543, 3.863	0.2264, 0.2313	0.2263, 0.2312	0.2265, 0.2313	0.2266, 0.2314	0.2504, 0.2575	0.2504, 0.2577	0.2503, 0.2576	0.2498, 0.2572
1000	3.503, 3.741	3.500, 3.749	3.494, 3.743	3.495, 3.743	0.2278, 0.2393	0.2269, 0.2387	0.2284, 0.2402	0.2294, 0.2419	0.2508, 0.2556	0.2508, 0.2555	0.2509, 0.2556	0.2509, 0.2556
5000	3.548, 3.651	3.547, 3.650	3.548, 3.652	3.549, 3.653	0.2259, 0.2277	0.2259, 0.2277	0.2260, 0.2277	0.2261, 0.2279	0.2512, 0.2533	0.2512, 0.2533	0.2512, 0.2533	0.2512, 0.2533
10000	3.633, 3.705	3.633, 3.705	3.633, 3.704	3.633, 3.704	0.2271, 0.2283	0.2271, 0.2282	0.2272, 0.2283	0.2272, 0.2283	0.2519, 0.2534	0.2519, 0.2534	0.2518, 0.2534	0.2518, 0.2534

Where: FONA =First Order Normal Approximation interval; BBI =Basic Bootstrap Interval, BPI = Bootstrap Percentile Interval; BCa = Bias-Corrected Accelerated interval

From Table 1, the confidence interval for the sample standard error of 10% trimmed mean exhibited a100% coverage of its estimated sample standard error in all four approaches and for all the values of B considered in this research work. This same confidence interval for the sample standard error of the 10% trimmed mean was equally getting more precise as the B the number of bootstraps increased. For instance, the estimated sample standard error for the 10% trimmed mean was s_{10} = 0.2266338; when B = 20, the bootstrap CI for the standard error returned (0.2051, 0.2343) for the normal approach CI, (0.2082, 0.2371) for the basic, (0.2071, 0.2360) for the percentile: and (0.2071, 0.2360) for the BCa method. As B increased to 5000, it returned (0.2259, 0.2277) for the normal approach, (0.2259, 0.2277) for the basic, (0.2260, 0.2277) for the percentile: and (0.2261, 0.2279) for the BCa method for the same estimated standard error. As B increased further to 10000, the CI for the sample standard error of the 10% trimmed mean returned a more precise interval of (0.2271, 0.2283) for the normal approach, (0.2271, 0.2282) for the basic, (0.2272, 0.2283) for the percentile:and (0.2272, 0.2283) for the BCa showing a shorter interval yet covering the estimate. It returned the best interval for our skewed data set.

The bootstrap CI for $s_{\bar{x}}$ (the sample standard error of the arithmetic mean)covered the estimated value only for B= 20 and 50. As B increased, to 100, the coverage seized for all four approaches and continued like that even up to B = 10000.

At B = 20, the CI for s_{20}(the standard error of the 20% trimmed mean) did not cover the estimated value. It rendered its best covering as B was increased to 50 and barely covered at B = 100, and 500. Beyond this point, the interval failed to cover the estimate again even for theBCa approach.

No significant difference or improvement in coverage was observed as a result of using a different approach even when B increased to 1000 and even got worse at B=5000 and B = 10000.

It appears that the effect of increasing B should be consideredonly after the right estimate has been chosen. For instance for skewed data like response data, the interest should be first on choosing the proper estimate like the alpha trimmed mean before one considers the approach to use.

No particular method appears to have specifically managed the effect of the skewness of data better than the other except for the parameter ten percent trimmed mean.

Even though having been adjusted for bias and expected to perform better even skewed data, the result showed that the α-trimmed mean out performed it for this set of response data.

We notice that the effect of increasing B was the same irrespective of the approach employed. The difference in its effect was only on the different estimates rather than on the approach employed.

Considering all we may conclude that for the set of data we considered, theconfidence interval for the standard error of the 10% trimmed mean returned the most accurate interval irrespective of the approach employed and the accuracy increased as B increased.

From Table 2, as B the number of bootstrapping increases, we notice that all four methods of calculating bootstrap CI considered returned shorter confidence intervals. Increasing B however did not improve coverage by any of the methods considered and even for all three statistics of interest.

Table 2: Confidence Interval Length Based on Table 1 above.

	Standard Error of the Mean ($s_{\bar{x}}$)				Standard Error of the 10% Trimmed Mean (s_{10})				Standard Error of the 20% Trimmed Mean (s_{20})			
B	FONA	BBI	BPI	BCa	FONA	BBI	BPI	BCa	FONA	BBI	BPI	BCa
20	1.894	1.8	1.8	1.579	0.0292	0.0289	0.0289	0.0289	0.0261	0.022	0.0219	0.0219
50	1.367	1.556	1.556	1.568	0.1161	0.1141	0.1141	0.1179	0.0135	0.0157	0.0157	0.0161
100	0.728	0.708	0.709	0.706	0.0092	0.0096	0.0096	0.0091	0.0191	0.0187	0.0186	0.0189
500	0.316	0.321	0.32	0.32	0.0049	0.0049	0.0048	0.0048	0.0071	0.0073	0.0073	0.0074
1000	0.238	0.249	0.249	0.248	0.0115	0.0118	0.0118	0.0125	0.0048	0.0047	0.0047	0.0047
5000	0.103	0.103	0.104	0.104	0.0018	0.0018	0.0017	0.0018	0.0021	0.0021	0.0021	0.0021
10000	0.072	0.072	0.071	0.071	0.0012	0.0011	0.0011	0.0011	0.0015	0.0015	0.0016	0.0016

Note: meaning of the abbreviations remains the same as in Table 1

We also observed from Table 2 that for CI of the standard error of the arithmetic mean, the BCa returned the shortest interval among all four approaches when B was as low as 20 returning a confidence interval length of 1.57. However, the confidence interval for the standard error of the arithmetic mean was longer than that for the α - trimmed mean for the same B=20. This could be because of the fact that the data is skewed. The BCa approach therefore: recommends itself for use when data is skewed and the arithmetic mean is the measure of central tendency. For this skewed data, the confidence interval for the sample α – trimmed mean performed better even when the B is as low as 20.

Increasing the number of bootstrapping B shortens the interval length but may not guarantee accuracy as it does not guarantee theinclusion ofa parameter within the interval exceptifthe right parameter is employed for the right set of data.

4. CONCLUSION AND RECCOMMENDATIONS.

4.1 Conclusion.

The 95% confidence interval for the standard error of the 10% trimmed mean returned accurate coverage of the true parameter by all four approaches applied.

In the light of the empirical study, we conclude that increasing B the number of bootstrap samples increases the chances of more precise and shorter confidence intervals rather than increasing the chances for coverage and accuracy of the interval.

One should not entirely neglect the shape of the data even when employing the bootstrap method of estimating the confidence interval no matter the approach used. This is to say that a better result is obtained when the right parameter is chosen for the right shape for instance, using the standard error of the alpha trimmed mean for a skewed data set, yields a better result irrespective of the approach.

4.2. Recommendations.

In general, increasing B produces shorter interval rather than increasing coverage/accuracy however, shorter interval is not a guarantee for better confidence interval unless the probability of coverage is first guaranteed. In the light of the study therefore, we make the following recommendations:

1. In estimating the bootstrap standard error, attention should be more on employing the appropriate measure of central tendency and its sample standard error rather than on arbitrarily increasing B the number of bootstrap samples.
2. When dealing with response data as in the study here, one should seriously consider its skewed nature and as such, the standard error of the alpha trimmed mean should be considered irrespective of the bootstrap method of estimating the CI.
3. Once the right tool is employed, increasing B will enhance the accuracy of the bootstrap CI.
4. With probability of coverage guaranteed, our findings tend to support the opinion of Efron, andTibshirani, (1986)[3] and later confirmed by that in estimating the bootstrap confidence interval, B = 1000 is enough for a well appreciable precise and short interval irrespective of the bootstrap approach employed. B = 500 can render a good short interval really.
5. If the standard error of the arithmetic mean is to be used, the BCa approach to estimating the confidence interval should considered first.

REFERENCES

1. Orloff, J. and Bloom, J. Bootstrap confidence intervals 18.05 class 24, Bootstrap con_dence intervals, Spring 2017 11Paraphrased from Dekking et al. A Modern Introduction to Probabilty and Statistics, Springer, 2005, 275.

2. Dixon, P.M. Bootstrap Resampling. In El-Shaarawi and A-H., Piegorsch, W.W. (Eds.-in-Chief). The Encyclopedia of Environmetrics (pp. 1-9). John Wiley & Sons, Ltd., 2006. https://doi.org/10.1002/9780470057339.vab028

3. Efron, B.& Tibshirani, R. Bootstrap Methods for Standard Errors, Confidence Intervals and Other Measures of Statistical Accuracy, *Statistical science*1986 1.(1) 54 -57; 1177013815

4. Davison, A. C. and Hinkley, D. V. Bootstrap Methods and their Application. Cambridge Uni-versity Press.1997. (statwww.epfl.ch/davison/BMA/)

5. Davidson, R., & MacKinnon, J. G. Bootstrap tests: How many bootstraps? *Econometric 2001. Reviews*, *19*(1), 55–68. https://doi.org/10.1080/07474930008800459

6. Wei, J.(2013) The Number of Bootstrap Replicates in Bootstrap Dickey-Fuller Unit Root Tests. Working Paper 2013:8 Available on www.statistics.uu.se E-mail: jianxin.wei@statistics.uu.se

7. Rousselet, G., Pernet, C. and Wilcox, R. A practical introduction to the bootstrap: a versatile method to make inferences by using data-driven simulations. 2019. [preprint].DOI:10.31234/osf.io/h8ft

8. Diez, D. (n.d.), "95% Confidence Intervals Under Variety of Methods," *https://docs.google.com/spreadsheets/d/ 1MNOCwOo7oPKrDB1FMwDzsYzvLoK-IBqoxhKrOsN1M2A/edit?pli=1#gid=0*.

9. Hesterberg, T. (2014), "What Teachers Should Know About the Bootstrap: Resampling in the Undergraduate Statistics Curriculum," available at *https://arxiv.org/pdf/1411.5279v1.pdf* .

10. Robert W. Hayden. Questionable Claims for Simple Versions of the Bootstrap, Journal of Statistics Education, 2019 27:3, 208-215, DOI: 10.1080/10691898.2019.1669507 Available on: https://doi.org/10.1080/10691898.2019.1669507

11. Heathcote, A., Popiel, S., and Mewhort, D. Analysis of Response Time Distributions: An Example Using the Stroop Task. Queen's University at KingstonKingston, Ontario, Canada Psychological Bulletin, 1991, 109(2), 340-347

12. Chernick, M. R. and LaBudde, R. A; *An Introduction to Bootstrap Methods with Applications to R*. John Wiley & Sons, Inc., Hoboken, New Jersey and Canada ,2011.

13. Lampert, E., Stahel, W., and Abbt, M. Lognormal Distributions Across the Sciences: Keys and Clues. BioScience, 2001, 51(5), 341-352

14. Rousselet, G., Pernet, C. and Wilcox, R. (2021).The Percentile Bootstrap: A Primer With Step-by-Step Instructions in R 2021. available on, https://doi.org/10.1177/2515245920911881

15. Rousselet, G. A., & Wilcox, R. R. *Reaction times and other skewed distributions: Problems with the mean and the median* 2019.[Preprint]. Available on: https://doi.org/10.31234/osf.io/3y54r

16. Davison, A. C., Hinkley, D. V. and Young G. A.(2003) Recent Developments in Bootstrap Methodology, *Statistical Science,* 2003, 18(2), 141–157, © Institute of Mathematical Statistics, 2003

17. Efron, B. Better Bootstrap Confidence Intervals. Tech. Rep. Stanford Univer. Dept. statist. 1984.

18. Hejase, A.J. and Hejase, H.J. Research Methods A Practical Approach for Business Students. 2nd Edition, Masadir Inc., Philadelphia; 2013.

19. Wilcox, R.R. Robust Estimation and Hypothesis Testing.2nd Edition. Elsevier, Academic Press. 2005. 30 Corporate Drive, Suite 400, Burlington, MA 01803, USA. ISBN: 0-12-751542-9.

20. Staudte, R.G. and Shearther, S. J. Robust Estimation and Testing. Wiley-Interscience Publication,John Wiley and Sons, Inc. 1990. New York ISBN 0-471-85547-2.

21. Madsen, R.W. and Moeschberger, M.L. Statistical Concepts with Applications to Business and Economics. Prentice Hall, Englewood Cliffs,1986. New Jersey 07632. ISBN 0-13-844846-9,

22. Efron, B., & Tibshirani, R. J. *An introduction to the bootstrap*.1993 New York: Chapman & Hall/CRC.

23. Canty, A. Package 'boot' 2021-05-03 09:04:02 UTC. Repository: CRAN
24. Larget, B. Chapter 3 R Bootstrap ExamplesFebruary 19, 2014 online

25. Efron B.Better bootstrap confidence intervals. *Journal of the American Statistical Association* 1987; 82:171200.

26. Lei, S., & Smith, M. R. Evaluation of several nonparametric bootstrap methods to estimate confidence intervals for software metrics. *IEEE Transactions on Software Engineering,* 2003; *29*, 996-1004.

27. Glasserman, P. Estimation and Confidence Interval. 403 Uris Hall, Columbia Bsiness School. Fall 2001, B6014: Managerial Statistics

28. Helwig, N. E. Bootstrap Confidence Intervals University of Minnesota (Twin Cities) Up