American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)

ISSN (Print) 2313-4410, ISSN (Online) 2313-4402

http://asrjetsjournal.org/

Bootstrap Methods and Reproducibility Probability

Sulafah Binhimd[a*], Bashair Almalki[b]

[a,b]*Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia*

[a]*Email: shamad@kau.edu.sa*

[b]*Email: balmalki0164@stu.kau.edu.sa*

Abstract

Bootstrap method is one of the resampling methods, it is a powerful and computer-based method. We will use two different bootstrap methods, Efron’s bootstrap and smoothed Efron’s bootstrap. They are resampling methods but in different manner and will be explained later in this paper. The reproducibility probability is the important topic will be discussed here and will be used to compare the two bootstrap methods. It is reflecting the stability of the results of hypothesis tests, it is the probability of obtaining the same decision when repeated hypothesis test.

Keywords: Bootstrap methods; Efron’s bootstrap method; Smoothed Efron’s bootstrap method; Reproducibility probability; Hypothesis test.

1. Introduction

One of the methods used for resampling is the bootstrap method. It is a powerful, computer-based method and does not require a lot of assumptions, the basic approach is based on resampling with replacement from the data set, to estimate statistics and their sampling distributions. Efron [1] explored that the bootstrap method uses Monte Carlo sampling to generate the empirical estimate of the statistics’ sampling distribution, see [2,3]. It is straightforward and easy in the application, so it has been used in the statistical inference such as confidence interval, regression and hypothesis tests. The bootstrap method has been developed over the years to produce different versions of bootstrap such as Bayesian bootstrap [4], smoothed Efron’s bootstrap and smoothed Bayesian bootstrap [5]. In this paper, we will use two versions of bootstrap, the Efron’s bootstrap and the smoothed Efron’s bootstrap. The smoothed bootstrap [5] smooths the Efron’s bootstrap by linear interpolation histospline, it is smoothing among the jump points of empirical distribution. It is depending on create n+1 intervals and then sample the observations from them.

----

* Corresponding author.

The smoothed Efron's bootstrap method can be used in several aspects such as linear regression [6], and prediction intervals [7]. The important topic that we will be discussing here too is the reproducibility probability (RP) [8]. The estimate of reproducibility probability is used with statistical hypothesis tests to reflect the stability of the results of tests, because it is concerned with the number of times we get the same first decision in the hypothesis test. During this paper, we will compare between two methods of bootstrap: Efron's bootstrap and smoothed Efron's bootstrap, using the estimate of RP. Section 2 will show an overview of the two methods of bootstrap, and how to estimate RP to compare them. In Section 3 we discuss the approach used in this paper with some results achieved. Section 4 show the conclusion of this study.

2. Materials and Methods

The reproducibility probability RP is important in application of statistical hypothesis tests, it is a quantity of randomness in a result [9]. In this paper, the RP is estimated using the percentage of cases that is reject (or not reject) the null hypothesis, we will focus on cases of rejection if the first decision using original sample is reject the null hypothesis, and focus on non-rejection cases if the first decision is non reject. The main gal of this paper is to compare between the two methods of bootstrap, which we will explain here, using the estimate of RP. The Efron's bootstrap method treat the original sample as a population. It is randomly drawing a large number of samples of size n from the original sample with replacement, the original sample here consists of independent observations. The smoothed Efron's bootstrap method uses different method to resample from the original sample, the n+1 intervals are created between n observations, then put uniformly distributed probabilities 1/(n+1) over each interval to sample n observations from them. Using the smoothed Efron's bootstrap method, the empirical distribution function is smoothed using linear interpolation histospline smoothing among the jump points. It spreads the probability 1/(n+1) uniformly over any interval between two values of observations. It is worth mentioning here that the n observations are real valued, one dimensional on a finite interval. For comparison the two bootstrap methods mentioned earlier will estimate RP by conducting hypothesis tests, Wilcoxon Signed Rank test and Sign test [10]. Wilcoxon Signed rank test is a nonparametric location test, it is a rank-based test, here we will deal with the case of one sample. The test statistics is the sum of the positive signed rank. One important assumption of this test is the distribution of the data is symmetric. While the sign test has no assumption about the shape of distribution. The sign test is also a nonparametric location test, it is test statistic is the total number of positive differences. To estimate RP, follow these steps:

1- Draw the original sample
2- Apply the hypothesis test and report the decision of test
3- Draw B bootstrap samples and apply the same test.
4- The estimation of RP is the ratio of times that have the same decision that the original sample has wither reject or non-reject the null hypothesis (in step 2).

The importance of RP is that it gives an indication of the stability the results of hypothesis tests. In the next section, we will apply the above steps with different cases of distributions, sample size and significance levels α.

3. Results and Discussion

Here we will use the two bootstrap methods we mentioned earlier to estimate the reproducibility probability RP. This study was applied in case of different sample sizes n=5, 20, 200, 500, 1000, 5000. The original sample was drawn from different distributions such as Uniform (0,1), Beta (0.7,0.7), Beta (2,10) and Beta (10,2). Figure 1 shows the value of RP when apply the Wilcoxon Signed Rank test with α= 0.01, 0.05, 0.1. We note that the smoothed Efron's bootstrap method has the smallest value of RP with n=5 and α=0.1, but in case of large samples there is no big difference between the two bootstrap methods. When α=0.05 the Efron's bootstrap method has the smallest value of RP with n=20, while in case of α=0.01 there is no difference between the two methods of bootstrap.

When apply the same test with another case, we show the results in Figure 2, which show to us that the smoothed Efron's bootstrap has the largest value of RP in most cases, except with n=5 and α=0.1. While the Efron's bootstrap method has small values of RP in case of small samples and with n=20 and α=0.05. With α=0.01 the difference between the two methods of bootstrap is simple. To apply the same approach with non-symmetric distributions we will use the Sign test. In Figure 3 the RP values of Sign test are shown, The Efron's bootstrap method are more stable, while the smoothed Efron's bootstrap method mostly get the largest value of RP at small samples. But the difference are very small here between the two methods and are negligible. In Figure 4, the smallest value of RP at n=20 with Efron's bootstrap method.

The largest value of RP appears at n=200 or more. At α=0.01 and n=20 the Efron's bootstrap method has the smallest value of RP. Given previous cases there is no difference between the two methods of bootstrap in case of large samples. But in case of small samples the smoothed Efron's bootstrap has more stable results.

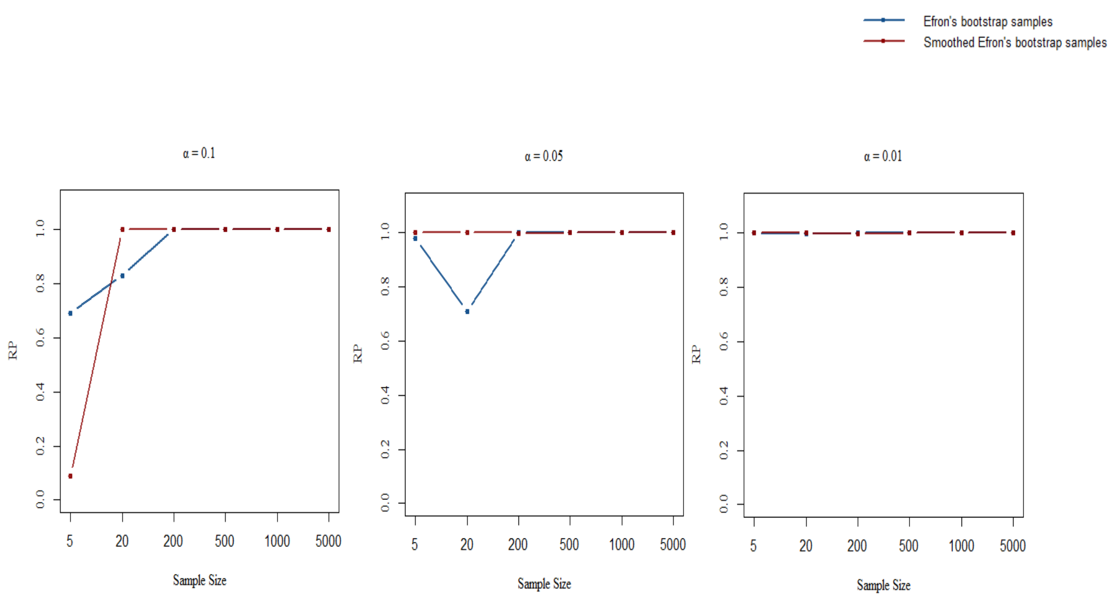

Figure 1: RP values if the original sample from Uniform (0,1)

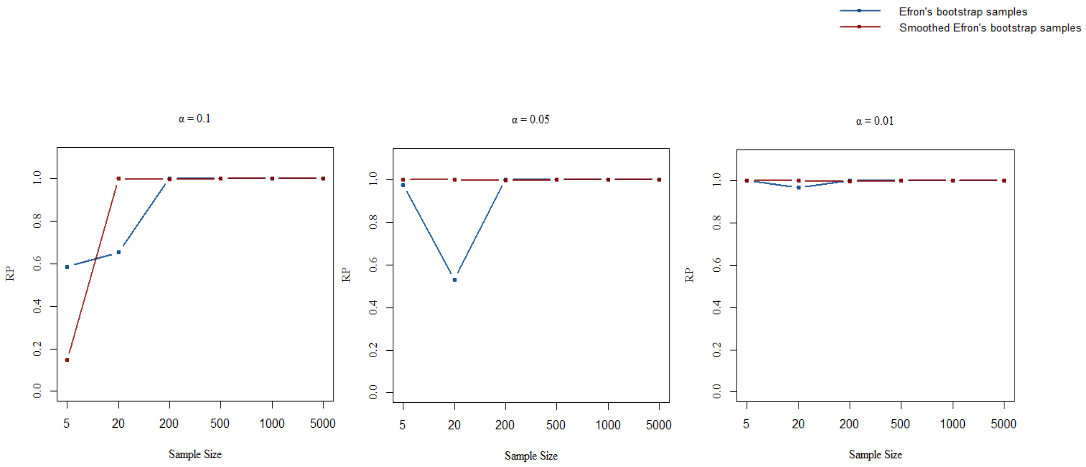

Figure 2: RP values if the original sample from Beta (0.7,0.7)

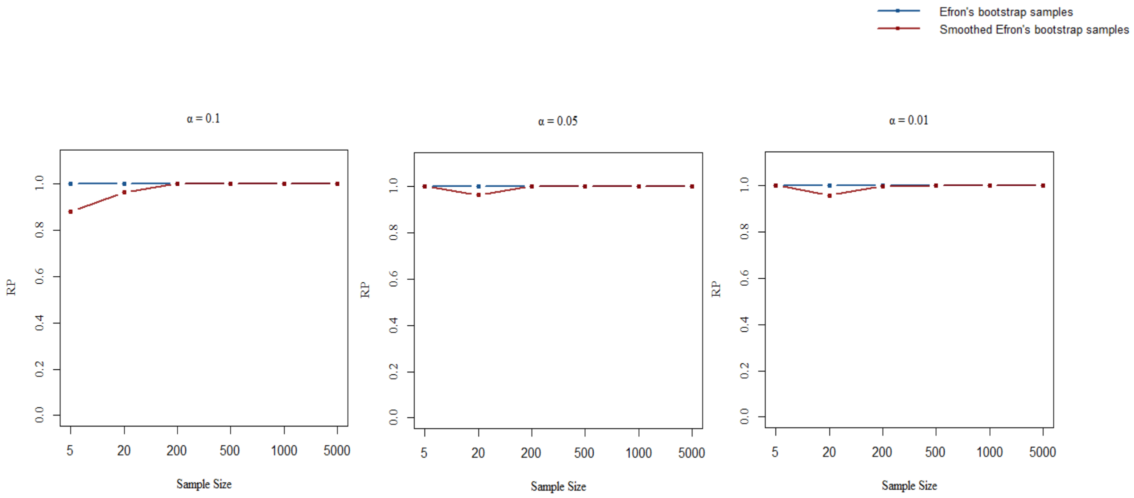

Figure 3: RP values if the original sample from Beta (2,10)

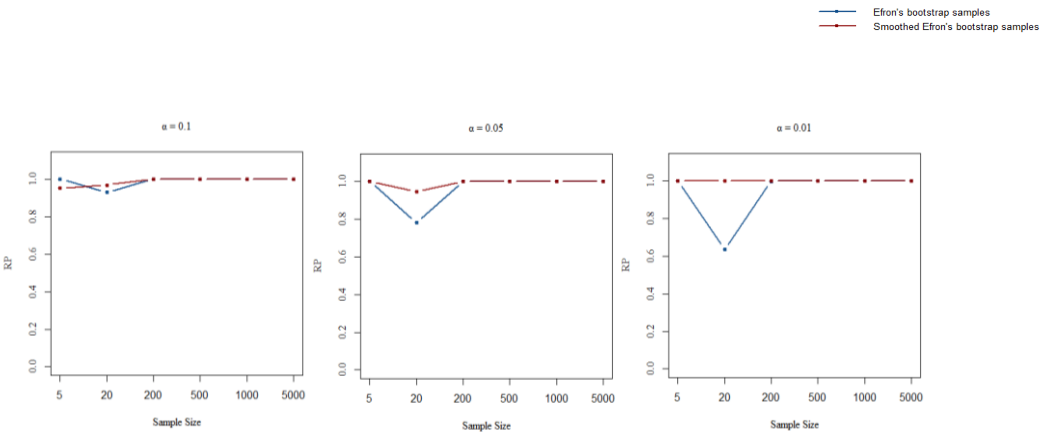

Figure 4: RP values if the original sample from Beta (10,2)

4. Conclusion

In this paper we compared the two methods of bootstrap, Efron's bootstrap and smoothed Efron's bootstrap, using the reproducibility probability. To do this we applied two of the nonparametric tests, Wilcoxon Signed Rank test and Sign test, to estimate the reproducibility probability in different cases. In general, we found that there is no difference between the two methods of bootstrap in the case of large samples, but the difference is shown in the case of small samples, as the values of RP using the smoothed Efron's bootstrap are more stable.

5. Recommendations

This paper compare between the two methods of bootstrapping, Efron's bootstrap and smoothed Efron's bootstrap, using the estimate of reproducibility probability. In this study a bootstrap sample size equal to the size of the original sample was used. To develop this study, sample sizes different from the original sample can be used to observe the effect on the study. Another way to extend this study is to generlize the smoothed Efron's bootstrap method to work with infinite interval instead of finite interval, this enables us to use other distributions for comparison. Another suggestion could include other types of bootstrap methods to become more general.

References

[1] B. Efron. "Bootstrap methods: Another look at the jackknife." The Annals of Statistics, vol. 7, pp.1-26, 1979.

[2] B. Efron, R.T. Tibshirani. "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy". Statistical Science, vol. 1, pp 54-75,1986.

[3] B. Efron, R.J. Tibshirani. An Introduction to the Bootstrap. Chapman and Hall, 1993.

[4] D.B. Rubin. "The Bayesian bootstrap". The Annals of Statistics, vol. 8, pp130-134, 1981.

[5] D.L. Banks. "Histospline smoothing the Bayesian bootstrap." Biometrika., vol. 4, pp. 673-684, 1988.

[6] S. Binhimd. "Bootstrap estimate of prediction error of simple linear regression models". American Scientific Research Journal for Engineering, Technology and Science, vol. 45, pp 274-279, 2018.

[7] S. Binhimd, B. Almalki. "Bootstrap prediction intervals". Applied Mathematical Science, vol. 12, pp 841-848, 2018.

[8] J. Shao, S. Chow. "Reproducibility probability in clinical trials". Statistics in Medicine, vol. 21, pp 1727-1742, 2002.

[9] L. De Capitani, D. De Martini. "Reproducibility probability estimation and RP-testing for some nonparametric tests". Entropy, vol. 18, pp 142, 2016.

[10] W.J. Conover. Practical nonparametric statistics. Wily, 1999.