CONSTRUCTION OF BOOTSTRAP CONFIDENCE INTERVALS FOR C_{NPK} IN CASE OF BIRNBAUM SAUNDERS DISTRIBUTION

Muhammad Kashif*[a, b], Muhammad Aslam[a], Ali Hussein AL-Marshadi[a], Khushnoor Khan[a], G. Srinivasa Rao[c], and Muhammad Imran Khan[b]

[a] Department of Statistics, Faculty of Sciences, King Abdul Aziz University, Jeddah (Saudi Arabia)

[b] Department of Mathematics and Statistics, University of Agriculture, Faisalabad, Pakistan.

[c] Department of Statistics, The University of Dodoma, Dodoma, PO. Box: 259, Tanzania

*Author for Correspondence; e-mail: mkashif@uaf.edu.pk

ABSTRACT: *The process capability index (PCI),C_{pk}, is commonly used in most of the manufacturing industries to measure process capability. An essential assumption while using C_{pk} is that the process should follow- a normal distribution. However, this may not be possible in certain cases; so the underlying process will follow non-normal distribution. For this reason usage of PCI based on non-normal distributions are gaining attention for researchers; and useful approach is to use the quantiles method. Traditionally, use of Clements method is preferred, whereas the application of Pearn and Chen method has limited usage. This paper aims to evaluate the C_{pk} for Birnbaum-Saunders (BS) distribution. Moreover, bootstrap confidence intervals for newly proposed index using Pearn and Chen method are also discussed. Three bootstrap confidence intervals: standard, percentile and biased corrected are compared based on average width and coverage probability.*

Keywords: Birnbaum-Saunders distribution, Non-parametric confidence intervals, Process Capability Index, Pearn and Chen method.

1.INTRODUCTION

One of the most important applications of statistical tool in manufacturing industries is to quantify process performance and method used for this purpose is process capability indices (PCIs). Among numerous PCI's, the most applicable index is C_{pk} [1-3]. After the development and application of these indices in industry, the primary focus of statisticians and other quality researchers has been on point estimation followed by the construction of confidence intervals for these indices[4]. Even the point estimator is a useful measure but confidence interval more helpful to think about the range of parameter of interest. [3]. As the PCIs are random variables and follow certain probability distributions; so confidence intervals plays a vital role for their correct interpretation [5]. The construction of confidence limits for PCI was started by [6]. Since then many techniques were developed to construct confidence limits for the PCIs. In the outset, most of the confidence intervals for the PCIs were constructed for a normally distributed process. But later on, the effort was diverted to develop estimation techniques those are free from the distributional assumptions because many processes are skewed or are heavy-tailed in practice. Recently efforts were made to construct confidence limits for PCIs and study their behaviors when the underlying process is non-normal. For this purpose a non-parametric statistical method called bootstrapping introduced by [7] is frequently used. The main attraction for the application of this approach is that it does not require the assumption of normality for constructing the confidence limits.

Birnbaum and Saunders [8] introduced the two parameters BS-distribution to model the physical behavior of fatigue crack growth under cyclic loading. If x ~ BS(γ, β), then probability density function f_x and the cumulative distribution function F_x are given below.

$$f(x,\gamma,\beta) = \frac{1}{2\gamma\beta\sqrt{2\pi}}\left[\left(\frac{\beta}{x}\right)^{\frac{1}{2}} + \left(\frac{\beta}{x}\right)^{\frac{3}{2}}\right] exp\left[-\frac{1}{2\gamma^2}\left(\frac{x}{\beta}+\frac{\beta}{x}-2\right)\right] \cdots (1)$$

$$F(x,\gamma,\beta) = \phi\left[\frac{1}{\gamma}\left\{\left(\frac{x}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{x}\right)^{\frac{1}{2}}\right\}\right] \cdots (2)$$

Where γ(>0) is the shape parameter and β>0 represents the scale parameter. The ϕ(.) is the standard normal CDF. The BS-distribution is unimodal, asymmetric and useful for describing skewed data. Nowadays, this distribution is frequently applied in almost all research areas due to its good properties [9]. In recent years, several authors [9-14] used the BS-distribution for quality related studies. Although the BS-distribution has frequently been used in many industrial processes, but its application to study the process capability index (C_{pk}) and its bootstrap confidence intervals are not common.

By exploring latest literature, it was found that many authors addressed the non-normality issues for different PCIs using different approaches [15-18]. There are five major approaches to the development of PCI for non-normal process [9]. Among these Clements method is mostly used, which uses the process quintiles [19]. The basic idea of Clements is to use the normal distribution property in which PCI is intended to yield only 0.27% of non-conforming products. In the aforementioned approach the variability (6σ) is substituted with $q_{0.9985}$-$q_{0.00135}$. The quantities, $q_{0.9985}$ and $q_{0.00135}$ are the 99.865th and 0.135th quantiles respectively. The new development in Clements approach is just to calculate one value i.e. $(q_{0.9985} - q_{0.001352}/2)$ instead of estimating two 3σ.

If a random variable X is normally distributed, i.e.X~N(μ,σ), then for given upper and lower specification limit the index C_{pk} is defined as

$$C_{pk} = min\left[\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right] \cdots \quad (3)$$

According to [4], the above index for non-normal process is defined as

$$C_{NPK} = min\left[\frac{2(USL - x(q_2))}{x(q_3)-x(q_1)}, \frac{2(x(q_2)-LSL)}{x(q_3)-x(q_1)}\right] \cdots \quad (4)$$

Where $x(\alpha)$ is the α^{th}quantile of the non-normal process i.e. $P(X < x_\alpha) = \alpha$ and $q_1 = 0.00135$, $q_2 = 0.5$ and $q_3 = 0.99865$. In this modified index, the center of the process is based on median because it considers a robust measure for skewed distribution.

Recently [9] introduced the BS-distribution in process capability related studies. They used the simulation approach to evaluate the performance of the proposed methodology. Authors of [9] have used BS-distribution to study the index *Cp*.

There are few published articles where construction of bootstrap confidence interval for C_{NPK} is applied to study the quality of industrial process using BS-distribution. This study is going to extend the usage of BS-distribution in construction of bootstrap confidence interval forC_{NPK}.

Thus, in this work, the main goal is to propose and develop the bootstrap confidence intervals using Pearn and Chen [4] method for C_{NPK} when the quality of interest follows BS-distribution. In the upcoming sections; relevant terminologies are introduced to describe PCI C_{NPK} and their bootstrap confidence intervals for BS-distribution. Then, Monte-Carlo simulation study is presented to demonstrate the above methodology in the next section. Finally, important conclusions have been drawn in the last section.

2.MATERIAL AND METHODS

This section explains the PCI, C_{pk}, when data follows a BS-distribution. The quantile function of BS-distribution given in Eq.(1) is defined as

$$\lambda(\Theta:\gamma,\beta) = \mathcal{S}(\Theta) = \frac{\beta}{4}\left[\gamma Z(\Theta) + \sqrt{\{\gamma Z(\Theta)\}^2+4}\right]^2 \quad \cdots (5)$$

Where $0 < \Theta < 1$ and $Z(\Theta)$ is the qth quantile of the standard normal distribution. Since $\mathcal{S}(0.5) = \beta$, so β is the median of the BS distribution. For the implementation of proposed index, the maximum likelihood (ML) estimator of shape and scale parameters using the BS-distribution are required. Using the log-likelihood function and its derivation with respect of $\gamma \& \beta$, the ML estimator of the shape parameter is given as

$$\hat{\gamma} = \left[\frac{s}{\hat{\beta}} + \frac{\hat{\beta}}{r} - 2\right]^{\frac{1}{2}} \cdots \quad (6)$$

Where $s = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample arithmetic mean and $r = \left[\frac{1}{n}\sum_{i=1}^{n} {x_i}^{-1}\right]^{-1}$is the sample harmonic mean. According to [8], the MLE of β denoted by $\hat{\beta}$ can be obtained as the unique positive root of the equation

$$\beta^2 - \beta[2r + K(\beta)] + r[S + K(\beta)] = 0 \cdots (7)$$

Since (7) is a non-linear equation in β, an iterative procedure has been used to compute $\hat{\beta}$. Using the invariance property of ML estimator, the ML estimator of C_{NPK}^{BS} is defined as

$$\hat{C}_{NPK}^{BS} = min\left[\frac{2(USL-\hat{\beta})}{\mathcal{S}(\Theta_3)-\mathcal{S}(\Theta_1)}, \frac{2(\hat{\beta}-LSL)}{\mathcal{S}(\Theta_3)-\mathcal{S}(\Theta_1)}\right] \cdots (8)$$

Where $\mathcal{S}(\Theta_1)$, $\hat{\beta}$ and $\mathcal{S}(\Theta_3)$ are the 0.135th, 50th and 99.865th quantiles of BS distribution respectively for specified shape and scale parameters.

2.1 The Method of Bootstrap:

The current study takes into account the construction of confidence intervals of the index, C_{NPK}^{BS} using three bootstrap techniques i.e. standard, percentile and bias corrected percentile, which is explained in this section.

Let $X_1, X_2, X_3, \cdots X_n$ be independent and identically distributed n random variables of interest. Then following steps are involved to explain the bootstrap procedure.

1. A bootstrap sample of size n is obtained from original sample by putting $1/n$ as mass at each point of X, and is denoted by ${X_1}^*, {X_2}^*, {X_3}^* \cdots {X_n}^*$.
2. Let ${X_m}^*$ where $1 \le m \le n$ be the mth bootstrap sample, then mth bootstrap estimator of C_{NPK}^{BS} is computed as

$$\hat{C}_{NPK}^{BS*} = min\left[\frac{2\left(USL-\mathcal{S}_{\Theta_2}(\hat{\theta}^m)\right)}{\mathcal{S}_{\Theta_3}(\hat{\theta}^m)-\mathcal{S}_{\Theta_1}(\hat{\theta}^m)}, \frac{2\left(\mathcal{S}_{\Theta_2}(\hat{\theta}^m)-LSL\right)}{\mathcal{S}_{\Theta_3}(\hat{\theta}^m)-\mathcal{S}_{\Theta_1}(\hat{\theta}^m)}\right] \cdots (9)$$

Where $\hat{\theta}^m$ is the mth estimator of parameter θ.

Since there are total n^n resamples. From these re-samples we calculate n^n values of $\hat{C}_{NPK}^{BS*}$. Each of these would be estimate of C_{NPK}^{BS}. The arrangement of the entire collection from smallest to largest, would constitute an empirical bootstrap distribution of $\hat{C}_{NPK}^{BS}$.

In this study, $B = 1000$ is assumed for bootstrap resamples.

2.1.1 Standard Bootstrap (SB) Confidence Interval:

From $B = 1000$ bootstrap estimates of $\hat{C}_{NPK}^{BS*}$, calculate the sample average and standard deviation as

$$\bar{C}_{NPK}^{*} = (1000)^{-1}\sum_{i-1}^{1000}\hat{C}_{NPK}^{BS*}(i)$$

$$S_{\hat{C}_{NPK}^{BS*}}^{*} = \sqrt{\left(\frac{1}{999}\right)\sum_{i=1}^{1000}(\hat{C}_{NPK}^{BS*}(i) - \bar{C}_{NPK}^{*})^2}$$

Thus the SB $(1-\alpha)100\%$ confidence interval is

$$CI_{SB} = \bar{C}_{NPK}^{*} \pm Z_{1-\frac{\alpha}{2}} S_{\hat{C}_{NPK}^{*}}^{*}$$

Where $Z_{1-\frac{\alpha}{2}}$ is obtained by using $\left(1-\frac{\alpha}{2}\right)^{th}$quantiles of the standard normal distribution.

2.1.2 Percentile Bootstrap (PB) Confidence Interval:

From the ordered collection of $\hat{C}_{NPK}^{BS*}(i)$, choose $100\left(\frac{\alpha}{2}\right)\%$ and the $100\left(1-\frac{\alpha}{2}\right)\%$ points as the end points of the confidence interval to give

$$CI_{PB} = \left(\hat{C}_{NPK}^{BS*}{}_{B\left(\frac{\alpha}{2}\right)}, \quad \hat{C}_{NPK}^{BS*}{}_{B\left(1-\frac{\alpha}{2}\right)}\right)$$

as the $(1-\alpha)100\%$ confidence interval of C_{NPK}. For a 95% confidence interval with $B = 1000$ is $CI_{PB} = \left(\hat{C}^{BS*}_{NPK(25)}, \quad \hat{C}^{BS*}_{NPK(750)}\right)$

2.1.3 Bias-Corrected Percentile Bootstrap (BCPB) confidence Interval:

Table 1: The estimated coverage probabilities and average widths of a 95% bootstrap confidence intervals of C_{Npk} for Birnbaum Saunders Distribution with $\beta = 1$

n	γ	true C_{Npk}	Average widths			Coverage probabilities		
			SB	PB	BCPB	SB	PB	BCPB
10	0.2500	1.2484	1.9419	1.8917	1.3858	0.9620	0.8126	0.8954
15	0.2500	1.2484	1.3428	1.3221	1.0854	0.9478	0.8516	0.9124
20	0.2500	1.2484	1.0725	1.0606	0.9167	0.9500	0.8796	0.9260
25	0.2500	1.2484	0.9185	0.9094	0.8113	0.9468	0.8922	0.9288
30	0.2500	1.2484	0.8118	0.8049	0.7325	0.9494	0.9048	0.9354
35	0.2500	1.2484	0.7390	0.7334	0.6765	0.9432	0.8998	0.9288
40	0.2500	1.2484	0.6789	0.6735	0.6288	0.9448	0.9108	0.9326
10	0.5000	0.5333	1.0084	0.9813	0.7114	0.9584	0.8082	0.8936
15	0.5000	0.5333	0.6981	0.6864	0.5585	0.9454	0.8494	0.9120
20	0.5000	0.5333	0.5575	0.5506	0.4723	0.9492	0.8774	0.9266
25	0.5000	0.5333	0.4773	0.4720	0.4183	0.9466	0.8902	0.9274
30	0.5000	0.5333	0.4215	0.4175	0.3777	0.9490	0.9082	0.9344
35	0.5000	0.5333	0.3836	0.3804	0.3491	0.9430	0.8984	0.9282
40	0.5000	0.5333	0.3523	0.3492	0.3245	0.9448	0.9096	0.9322
10	0.7500	0.2953	0.6770	0.6564	0.4618	0.9612	0.8024	0.8920
15	0.7500	0.2953	0.4626	0.4535	0.3604	0.9504	0.8466	0.9104
20	0.7500	0.2953	0.3661	0.3608	0.3035	0.9528	0.8736	0.9248
25	0.7500	0.2953	0.3117	0.3077	0.2682	0.9506	0.8872	0.9286
30	0.7500	0.2953	0.2740	0.2709	0.2417	0.9512	0.8992	0.9348
35	0.7500	0.2953	0.2487	0.2462	0.2233	0.9462	0.8970	0.9278
40	0.7500	0.2953	0.2277	0.2254	0.2072	0.9476	0.9074	0.9318
10	1.0000	0.1849	0.4981	0.4810	0.3265	0.9686	0.7944	0.8900
15	1.0000	0.1849	0.3332	0.3257	0.2520	0.9574	0.8440	0.9100
20	1.0000	0.1849	0.2603	0.2559	0.2108	0.9600	0.8704	0.9226
25	1.0000	0.1849	0.2198	0.2166	0.1856	0.9550	0.8852	0.9288
30	1.0000	0.1849	0.1920	0.1895	0.1667	0.9536	0.8972	0.9338
35	1.0000	0.1849	0.1737	0.1717	0.1538	0.9510	0.8954	0.9284
40	1.0000	0.1849	0.1584	0.1566	0.1424	0.9532	0.9062	0.9320
10	1.2500	0.1255	0.3850	0.3702	0.2427	0.9786	0.7896	0.8894
15	1.2500	0.1255	0.2514	0.2451	0.1850	0.9642	0.8396	0.9090
20	1.2500	0.1255	0.1937	0.1901	0.1536	0.9648	0.8660	0.9214
25	1.2500	0.1255	0.1623	0.1596	0.1347	0.9594	0.8842	0.9272
30	1.2500	0.1255	0.1408	0.1388	0.1206	0.9586	0.8964	0.9326
35	1.2500	0.1255	0.1269	0.1254	0.1111	0.9548	0.8930	0.9274
40	1.2500	0.1255	0.1154	0.1140	0.1027	0.9566	0.9058	0.9314
10	1.5000	0.0903	0.3077	0.2945	0.1871	0.9848	0.7852	0.8864
15	1.5000	0.0903	0.1961	0.1906	0.1409	0.9734	0.8362	0.9088
20	1.5000	0.0903	0.1491	0.1460	0.1162	0.9704	0.8644	0.9216
25	1.5000	0.0903	0.1239	0.1218	0.1015	0.9638	0.8822	0.9276
30	1.5000	0.0903	0.1070	0.1054	0.0906	0.9620	0.8942	0.9324
35	1.5000	0.0903	0.0962	0.0949	0.0834	0.9582	0.8918	0.9280
40	1.5000	0.0903	0.0871	0.0860	0.0769	0.9610	0.9044	0.9316

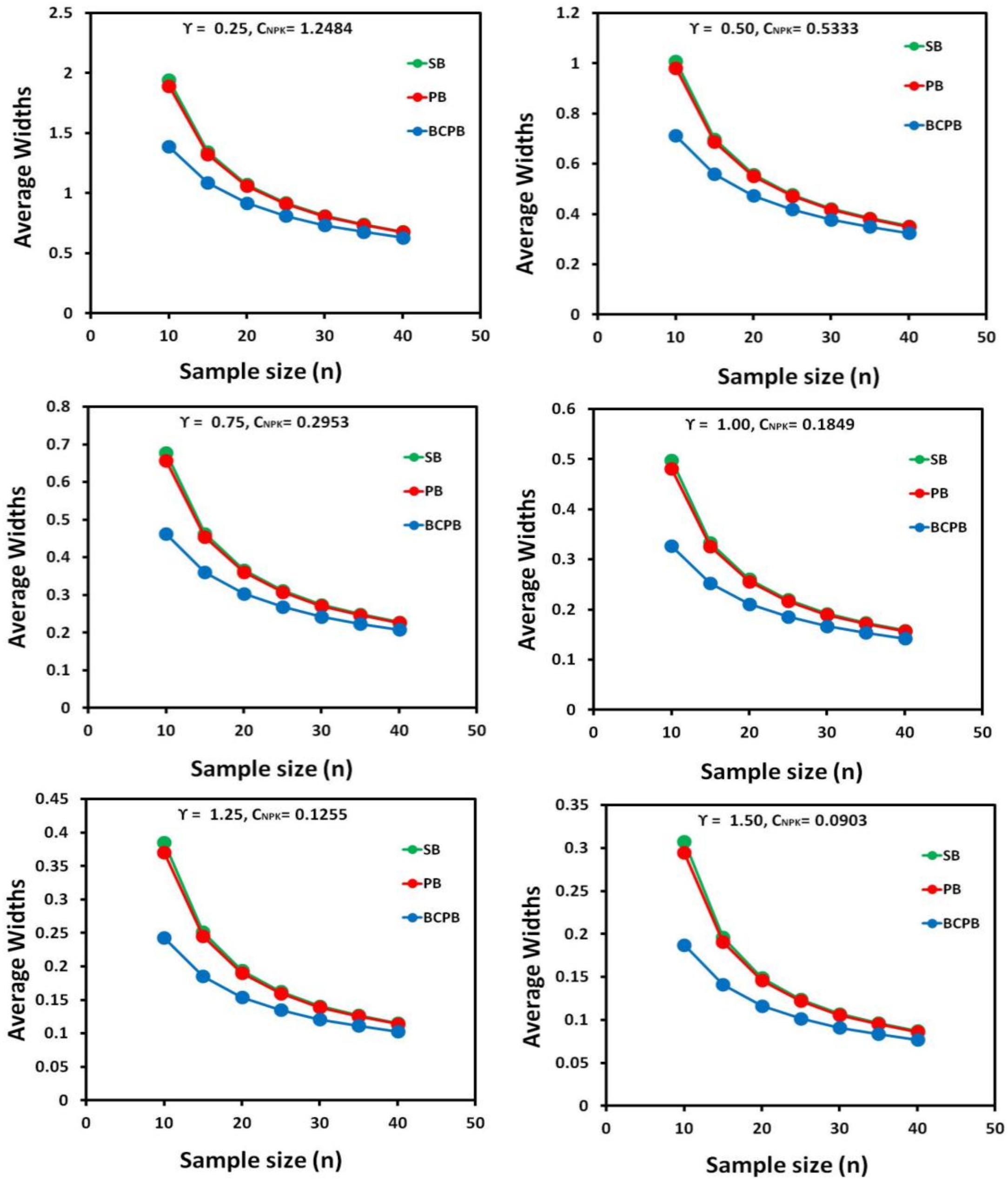

Figure 1. Comparison of average width of $\hat{C}_{NPK}^{BS}$ using SB,PB and BCPB method using different sample sizes and shape parameter.

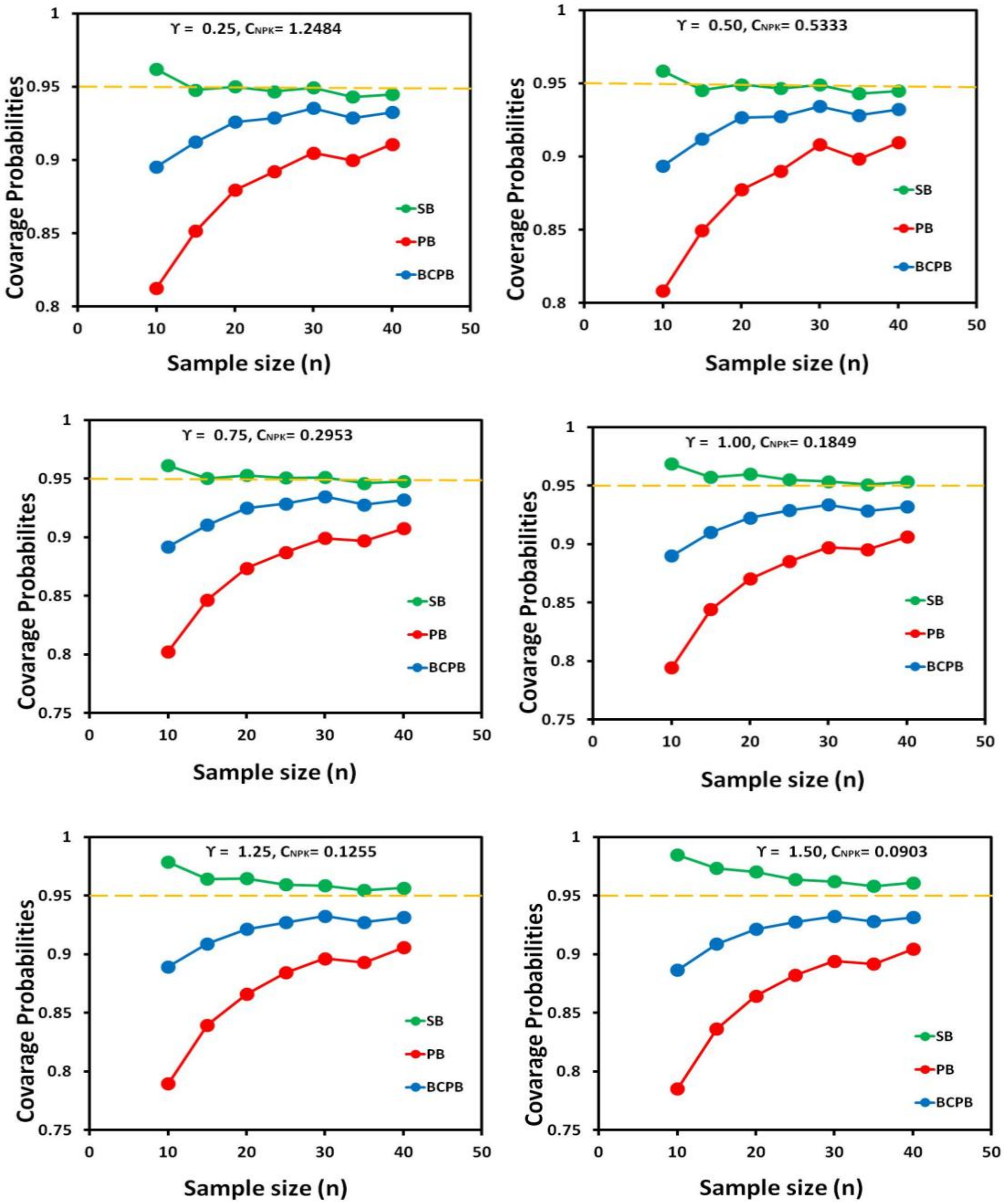

Figure 2. Comparison of coverage probabilities of $\hat{C}_{NPK}^{BS}$ using SB,PB and BCPB method using different sample sizes and shape parameter.

This method corrects the potential bias. Bias is generated because the bootstrap distribution is based on a sample from the complete bootstrap distribution and may be shifted higher or lower that would be expected. The calculation of this method is based on the following steps.

i. Using the (ordered) distribution of $\hat{C}_{NPK}^{BS*}(i)$, compute the probability

$$p_0 = pr\left(\hat{C}_{NPK}^{BS*} \leq \hat{C}_{NPK}^{BS}\right)$$

ii. Let $\emptyset$ and $\emptyset^{-1}$ represents the cumulative and inverse cumulative distribution functions of standard normal variable z, then calculate

$$z_0 = \emptyset^{-1}(p_0)$$

iii. The percentiles of the ordered distribution of $\hat{C}_{NPK}^{*}$ is obtained as

$$P_L = \emptyset\left(2z_0 + z_{\frac{\alpha}{2}}\right)$$

$$P_U = \emptyset\left(2z_0 + z_{1-\frac{\alpha}{2}}\right)$$

Finally the BCPB confidence interval is given as

$$CI_{BCPB} = \left(\hat{C}_{NPK(P_LB)}^{BS*}, \quad \hat{C}_{NPK(P_UB)}^{BS*}\right)$$

3.RESULTS AND DISCUSSION

This section evaluates the performance of the three non-parametric confidence intervals i.e. SB, PB and BCPB. Different data sets are generated using the PDF of BS-distribution given in Eq (1) with scale parameter, $\beta = 1$ and different shape parameter, $\gamma = 0.25, 0.50, 0.75, 1.00, 1.25$ and 1.50. For each combination of scale and shape parameter, a single sample of size $n = 10,15,20,25,30,35\ and\ 40$ was drawn. Then $B = 1000$ bootstrap resample were drawn from single samples of same sizes to estimate the coverage probabilities and average width. The 95% bootstrap confidence limits and their coverage probabilities were constructed by each of the three methods for index C_{NPK}. The complete simulations were run using the R-software. Table 1 exhibits the average width and coverage probabilities of three bootstrap methods. It is observed that estimated coverage probabilities of both SB and BCPB tend to increase towards nominal confidence levels for increased sample size. Same trend is observed in case of PB but with a slower rate. Table 1 also shows that CI_{SB} had a greater coverage probabilities and large average widths as compare to the other two intervals. The coverage probabilities show the following order $SB > BCPB > PB$. While in case of average width following order is observed $BCPB < PB < SB$.

The average widths of all confidence intervals reduce by increasing sample sizes. The performance of BCBP method is better as compared to the other two methods. Figure 1 is the graphical representation of average widths of three confidence intervals for different sample sizes as reported in Table 1. It is observed that smaller sample sizes have larger width. The width decreases rapidly as the sample size increases for all three methods. Sample size affects the interval's width in all cases. It can be observed that the behavior of SB and PB method is almost similar and BCBP method performs better than those of SB and PB. Figure 2 shows the coverage probability of each method under different sample sizes. The coverage probability of SB and PB method attain the standard nominal values when the sample size increases. It can be observed that the coverage rate ranges between 89 % to 94% for SB and BCPB except for PB which has lower estimate for all sample sizes.

4.Conclusions

Three bootstrap confidence intervals and their coverage probabilities are calculated for process capability index C_{NPK} using BS-distribution. The results indicate that both shape parameter and sample sizes affect the width and coverage probabilities of confidence interval. The width decreases as sample size increases. On the other hand, coverage probabilities reach near to nominal levels with the increase of sample size. Based on coverage probabilities and average widths BCPB method is recommended.

REFERENCES

1. Kane, V.E., *Process capability indices.* Journal of quality technology, 1986. **18**(1): p. 41-52.
2. Zhang, J., *Conditional confidence intervals of process capability indices following rejection of preliminary tests*. 2010, The University of Texas at Arlington.
3. Panichkitkosolkul, W. (2014). Confidence Interval for the Process Capability Index pC Based on the Bootstrap-t Confidence Interval for the Standard Deviation. *Metodološki zvezki, 11*(2), 79-92.
4. Pearn, W. and K. Chen, *Capability indices for non-normal distributions with an application in electrolytic capacitor manufacturing.* Microelectronics Reliability, 1997. **37**(12): p. 1853-1858.
5. Balamurali, S., *Bootstrap confidence limits for the process capability index Cpmk.* International Journal of Quality Engineering and Technology, 2012. **3**(1): p. 79-90.
6. Chou, Y.-M., D. Owen, and S. BORREGO A, *Lower confidence limits on process capability indices.* Journal of Quality Technology, 1990. **22**(3): p. 223-229.
7. Efron, B., *Bootstrap methods: another look at the jackknife.* The annals of Statistics, 1979: p. 1-26.
8. Birnbaum, Z. and S.C. Saunders, *A new family of life distributions.* Journal of applied probability, 1969: p. 319-327.
9. Leiva, V., C. Marchant, H. Saulo, M. Aslam and F. Rojas (2014). "Capability indices for Birnbaum–Saunders processes applied to electronic and food industries." Journal of Applied Statistics 41(9): 1881-1902.
10. Balakrishnan, N., V. Leiva, and J. López, *Acceptance sampling plans from truncated life tests based on the generalized Birnbaum–Saunders distribution.* Communications in Statistics—Simulation and Computation, 2007. **36**(3): p. 643-656.
11. Lio, Y. and C. Park, *A bootstrap control chart for Birnbaum–Saunders percentiles.* Quality and Reliability Engineering International, 2008. **24**(5): p. 585-600.
12. Lio, Y., T.-R. Tsai, and S.-J. Wu, *Acceptance sampling plans from truncated life tests based on the Birnbaum–Saunders distribution for percentiles.* Communications in Statistics-Simulation and Computation, 2009. **39**(1): p. 119-136.
13. Aslam, M., C.-H. Jun, and M. Ahmad, *New acceptance sampling plans based on life tests for Birnbaum–Saunders distributions.* Journal of Statistical Computation and Simulation, 2011. **81**(4): p. 461-470.
14. Leiva, V. Gerson,S. Enrique,C. and Guillermo,C. *New control charts based on the Birnbaum-Saunders distribution and their implementation.* Revista Colombiana de Estadística, 2012. **34**(1): p. 147-176.
15. Sennaroglu, B. and O. Senvar, *Performance comparison of box-cox transformation and weighted variance methods with weibull distribution. journal of aeronautics and space technologies*, 2015. 8(2): p. 49-55.
16. Kashif,M. Aslam,M. Al- Marshadi, A.H, and C.-H. Jun. *Capability Indices for Non-Normal Distribution Using Gini's Mean Difference as Measure of Variability*. IEEE Access, 2016. 4: p. 7322-7330.
17. Panichkitkosolkul, W., *Confidence Intervals for the Process Capability Index Cp Based on Confidence Intervals for Variance under Non-Normality*. Malaysian Journal of Mathematical Sciences, 2016. 10(1): p. 101–115.
18. Senvar, O. and B. Sennaroglu, *Comparing performances of clements, box-cox, Johnson methods with weibull distributions for assessing process capability*. Journal of Industrial Engineering and Management, 2016. 9(3): p. 634.

19. Clements, J.A., *Process capability calculations for non-normal distributions.* Quality progress, 1989. **22**(9): p. 95-97.