ORIGINAL RESEARCH

Bootstrap confidence intervals of C_{Npk} for type-II generalized log-logistic distribution

Srinivasa Rao Gadde[1] · K. Rosaiah[2] · SVSVSV Prasad[2,3]

Abstract

This paper deals with construction of confidence intervals for process capability index using bootstrap method (proposed by Chen and Pearn in Qual Reliab Eng Int 13(6):355–360, 1997) by applying simulation technique. It is assumed that the quality characteristic follows type-II generalized log-logistic distribution introduced by Rosaiah et al. in Int J Agric Stat Sci 4(2):283–292, (2008). Discussed different bootstrap confidence intervals for process capability index. Maximum likelihood method is considered for obtaining the estimators of the parameter. Monte Carlo simulation technique is applied to find out the coverage probabilities and average widths of the bootstrap confidence intervals. The results are illustrated with real data sets.

Keywords Process capability index · Bootstrap confidence interval · Maximum likelihood estimate · Type-II generalized log-logistic distribution

Introduction

In the present era, the term "Process Capability Indices" finds frequent space in the statistical quality control literature. If one may keen to study whether the ongoing production process is moving according to the predefined specifications or not, process capability index is the right technique to be chosen as it will help in monitoring and analyzing quality process and productivity. As we know, statistical quality control refers to the use of statistical technique in monitoring and maintaining the standards of products and services. A quality process evaluation procedure such as process capability analysis helps the manufacturer to achieve consumer quality expectations. It is an effective measure to gauge the quality of production process. Process capability analysis is used to determine whether the process capability of a supplier conforms to a customer's specifications, by applying an expression called the process capability index (PCI), to a controlled process. PCI is one such tool to measure the quality process at given specifications. As this method is simple and transparent and underlying assumptions are not complicated compared with conventional methods, PCI methods got popularized. The first PCI was developed by Juran (1974); later, different PCIs found their origin when the underlying distribution is normal, viz. Vännman (1995), Aslam et al. (2013), etc. As pointed by Kane (1986) and Gunter (1989a, b, c, d), the quality characteristic may not be normal in many occasions and assuming normality in such cases could lead to inaccurate and unreliable results.

Clements (1989) introduced two PCIs, C_p and C_{pk}, for non-normal data by relaxing the normality assumption and brought the concept of quantile-based PCIs. Later, it was developed by Vännman (1995). In a similar way, Kane (1986) developed PCI based on non-normal assumption. Distinguished statisticians made their efforts in developing different PCI methods, viz. Chan et al. (1988), Pearn and Chen (1995), Chen and Pearn (1997), Wood (2005), Chen et al. (2008), Wu and Liang (2010), Perakis (2010) and Kashif et al. (2017).

Peng (2010a, b) developed parametric lower confidence limits of quantile-based PCIs and also studied PCIs for

✉ Srinivasa Rao Gadde
  gaddesrao@yahoo.com

  K. Rosaiah
  rosaiah1959@gmail.com

  SVSVSV Prasad
  svsvsvprasad@gmail.com

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

[2] Department of Statistics, Acharya Nagarjuna University, Guntur 522 007, India

[3] Reserve Bank of India, Nagpur 440001, India

processes with skewed distributions. Similar developments can be seen from Kantam et al. (2010) and Wararit and Somchit (2012) for half-logistic distribution, and from Rao et al. (2015) for inverse Rayleigh and log-logistic distributions. The main aim behind the development of PCIs is to give indication on the quality process whether it is moving in line with the predefined standards. These standards can be determined by setting lower specification limit (*L*) and upper specification limit (*U*). In the traditional approach, it is assumed that the quality process is normally distributed, and then the PCI, C_{pk}, is given by

$$C_{pk} = \text{Min}\left\{ \frac{\text{USL} - \mu}{3\sigma}, \frac{\mu - \text{LSL}}{3\sigma} \right\}. \tag{1}$$

The sample mean ($\bar{x}$) and standard deviation (*s*) derived from a random sample of size $n\left(X_1, X_2, \ldots, X_n\right)$ are to be used to estimate the unknown parameters μ and σ; hence,

$$\tilde{C}_{pk} = \text{Min}\left\{ \frac{\text{USL} - \bar{x}}{3s}, \frac{\bar{x} - \text{LSL}}{3s} \right\}. \tag{2}$$

Clements (1989) suggested that if the process characteristic is drawn from a non-normal distribution, PCI C_{pk} can be constructed for any distribution as

$$C_{pk} = \text{Min}\left\{ \frac{\text{USL} - M}{U_p - M}, \frac{M - \text{LSL}}{M - L_p} \right\}. \tag{3}$$

U_p, L_p and *M* are, respectively, the 99.865th, 0.135th and 50th percentiles of the concerned distribution. Another method proposed by Chen and Pearn (1997), when the underlying process is from a non-normal distribution, is

$$C_{Np}(u;v) = \frac{d - u\left|\xi_{p2} - m\right|}{3\sqrt{\left(\left(\xi_{p3} - \xi_{p1}\right)/6\right)^2 + v\left(\xi_{p2} - T\right)^2}}, \tag{4}$$

where ξ_q is the *q*th quantile, i.e., $P\left(X < \xi_q\right) = q$, $p_1 = 0.00135, p_2 = 0.5, p_3 = 0.99865, d = (\text{USL} - \text{LSL})/2$, $m = (\text{USL} + \text{LSL})/2$ and *T* is the target value; from (4), we have

$$C_{Np}(0, 0) = C_{Np}, C_{Np}(0, 1) = C_{Npm}, C_{Np}(1, 0) = C_{Npk}$$
$$\text{and} \quad C_{Np}(1, 1) = C_{Npmk}.$$

As described above, many PCI methods are developed so far, and few of the widely used PCIs are C_p and C_{pk} developed by Kane (1986). In this paper, we proposed the popular process capability index C_{Npk} when the quality process follows TGLLD. The next part of the article is prepared in the following way. Introduction of TGLLD and the estimation of parameters using ML method are given in "Type-II generalized log-logistic distribution" section. In "Bootstrap confidence intervals" section, the bootstrap confidence intervals are determined for the PCIs proposed by Chen and Pearn (1997). In "Simulation study" section, simulated results for small sample comparison are tabulated. Finally, the benefit of PCIs so developed for TGLLD is demonstrated with an example in "Illustrative example" section.

Type-II generalized log-logistic distribution

Log-logistic distribution (LLD) has proven its importance in quality control. Different authors developed properties and types of acceptance sampling plans for LLD. The cumulative distribution function (CDF) of the log-logistic distribution (LLD) is

$$F(t;\sigma, \theta) = \frac{(t/\sigma)^\lambda}{\left[1 + (t/\sigma)^\lambda\right]}; t > 0, \sigma > 0, \lambda > 1, \tag{5}$$

where σ is the scale parameter and λ is the shape parameter.

The practical pertinence of generalized log-logistic distribution (GLLD) in diverse sectors attracted various authors to put their attention in developing some extensions for effective and wide use of log-logistic distribution, viz. Rosaiah et al. (2006, 2007). One such extension to this distribution is named as type-II generalized log-logistic distribution (TGLLD) introduced by Rosaiah et al. (2008); its cumulative distribution function (CDF) is

$$F(t;\lambda, \theta, \sigma) = 1 - \left[1 + (t/\sigma)^\lambda\right]^{-\theta}; \quad t > 0, \lambda > 1, \theta > 0, \sigma > 0. \tag{6}$$

It may be noted that the distribution given in (6) is defined through the reliability-oriented generalization of log-logistic distribution. In short, we call this as the type-II generalized log-logistic distribution [type-I generalized (exponentiated) log-logistic distribution is given by Rosaiah et al. (2006)]. The corresponding probability density function (PDF) is given by

$$f(t;\lambda, \theta, \sigma) = \frac{\lambda\theta}{\sigma} \frac{(t/\sigma)^{\lambda-1}}{\left[1 + (t/\sigma)^\lambda\right]^{\theta+1}}; \quad t > 0, \lambda > 1, \theta > 0, \sigma > 0, \tag{7}$$

where σ is the scale parameter, and λ and θ are shape parameters. The three-parameter TGLLD will be denoted by TGLLD $(\sigma, \theta, \lambda)$. If $\theta = 1$, then Eq. (7) becomes log-logistic distribution, and if $\lambda = 1$, then TGLLD becomes Pareto type-II distribution. Since log-logistic distribution is also a survival model as exemplified by many authors in the past, a series system of independent components with common log-logistic lifetime distribution for each component, we are motivated to study some inferential aspects of the distribution of such a series system. As not much work is reported

about the study of such a model, we made an attempt to take up some theoretical, applied inferential problems with respect to type-II generalized log-logistic model. The model provided more accurate results, especially when the data were examined for quality characteristics. Rao et al. (2012a, b) developed the reliability test plans for this distribution. The reliability function and hazard (failure rate) function of type-II generalized log-logistic distribution are, respectively, given by

$$R(t) = \left[1 + (t/\sigma)^{\lambda} \right]^{-\theta} \tag{8}$$

$$h(t) = \frac{\lambda \theta (t/\sigma)^{\lambda-1}}{\left[1 + (t/\sigma)^{\lambda} \right]}; \quad t > 0, \lambda > 1, \theta > 0, \sigma > 0. \tag{9}$$

Let $t_1, t_2, \ldots, t_n$ be a random sample of size n drawn from TGLLD $(T; \theta, \lambda, \sigma)$, then likelihood function L of the sample is

$$L = \prod_{i=1}^{n} f(t_i; \theta, \lambda, \sigma) = \frac{\lambda^n \theta^n}{\sigma^{n\lambda}} \prod_{i=1}^{n} \frac{(t_i)^{\lambda-1}}{\left[1 + \left(t_i/\sigma \right)^{\lambda} \right]^{\theta+1}}. \tag{10}$$

The log-likelihood function is

$$\log L = n \log \lambda + n \log \theta - n\lambda \log \sigma + (\lambda - 1)$$
$$\times \sum_{i=1}^{n} \log t_i - (\theta + 1) \sum_{i=1}^{n} \log \left[1 + \left(t_i/\sigma \right)^{\lambda} \right]. \tag{11}$$

The log-likelihood equations to obtain MLEs of θ, λ and σ are obtained as

$$\frac{\partial \log L}{\partial \sigma} = 0 \Rightarrow \frac{-n\lambda}{\sigma} + \frac{\lambda(\theta+1)}{\sigma} \sum_{i=1}^{n} \frac{\left(t_i/\sigma \right)^{\lambda}}{\left[1 + \left(t_i/\sigma \right)^{\lambda} \right]} = 0 \tag{12}$$

$$\frac{\partial \log L}{\partial \lambda} = 0 \Rightarrow \frac{n}{\lambda} - n \log \sigma + \sum_{i=1}^{n} \log t_i - (\theta + 1)$$
$$\times \sum_{i=1}^{n} \frac{1}{\left[1 + \left(t_i/\sigma \right)^{\lambda} \right]} \left(t_i/\sigma \right)^{\lambda} \log \left(t_i/\sigma \right) = 0 \tag{13}$$

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^{n} \log \left[1 + \left(t_i/\sigma \right)^{\lambda} \right]$$

MLE of θ is $\hat{\theta} = \dfrac{n}{\sum\limits_{i=1}^{n} \log \left[1 + \left(t_i/\sigma \right)^{\lambda} \right]}.$ (14)

Let the parametric function of TGLLD be represented by $\tau = (\theta, \lambda, \sigma)$; by recalling the invariance property of MLE, $\hat{\xi}_q = \xi_q(\hat{\tau})$ becomes the maximum likelihood estimator (MLE) of quantile ξ_q, where $\hat{\tau} = \left(\hat{\theta}, \hat{\lambda}, \hat{\sigma} \right)$ be the MLE of $\tau = (\theta, \lambda, \sigma)$. Hence, we consider that the MLE of the proposed PCI with parametric function $\hat{\tau} = \left(\hat{\theta}, \hat{\lambda}, \hat{\sigma} \right)$ is

$$\hat{C}_{Npk}(\tau) = C_{Npk}(\hat{\tau}) = \frac{\min \left(\text{USL} - \xi_{p_2}(\hat{\tau}), \xi_{p_2}(\hat{\tau}) - \text{LSL} \right)}{\left(\left(\xi_{p_3}(\hat{\tau}) - \xi_{p_1}(\hat{\tau}) \right)/2 \right)}.$$
(15)

Therefore, $C_{Npk}(\hat{\tau})$ is a real-valued function of quantiles ξ_{p_1}, ξ_{p_2} and ξ_{p_3}.

The qth quantile of TGLLD with parameters $\tau = (\theta, \lambda, \sigma)$ is given by $\xi_q = \sigma \left[(1-q)^{-1/\theta} - 1 \right]^{1/\lambda}$. Simulation technique is considered for small sample comparison, as it may be difficult to find out mathematical form of sampling distribution of $\hat{\tau}$.

Bootstrap confidence intervals

Bootstrap sampling is a method of drawing samples (with replacement) from the underlying probability distribution. Efron (1982) introduced a computationally intensive method of estimation called *Bootstrap*, a technique of computer-based simulation for estimating the parameters under consideration. Estimates of PCI are determined through bootstrap confidence intervals. As stated by Efron and Tibshirani (1993), for obtaining reasonably accurate confidence interval estimates, a minimum of 1000 bootstrap resamples may be considered. Among many other methods, the types of bootstrap confidence interval developed by Efron and Tibshirani are considered under study, viz. the standard bootstrap (SB) confidence interval, the percentile bootstrap (PB) confidence interval and the bias-corrected percentile bootstrap (BCBP) confidence interval.

Let $t_1, t_2, \ldots, t_n$ be a random sample of size n drawn from a quality process following TGLLD; then, $t_1^*, t_2^*, \ldots, t_n^*$ is a bootstrap sample of size n drawn with replacement from the original sample. Using this bootstrap sample, the bootstrap $\hat{C}_{Npk}$, denoted by $\hat{C}_{Npk}^*$, can be obtained. For B bootstrap samples, we can obtain B bootstrap $\hat{C}_{Npk}'$s and arrange them in ascending order, i.e., $\hat{C}_{Npk}^*(1), \hat{C}_{Npk}^*(2), \ldots, \hat{C}_{Npk}^*(B)$, which forms an empirical bootstrap distribution of $\hat{C}_{Npk}$. Here, we take $B = 10{,}000$ bootstrap samples.

Standard bootstrap (SB) confidence interval

Here, the PCI estimate under study is $\hat{C}_{Npk}$; then, *jth* bootstrap sample estimator of $\hat{C}_{Npk}$ is

$$\hat{C}_{Npk}^{*(j)} = \frac{\min\left(\text{USL} - \xi_{p_2}(\hat{\tau}^{(j)}), \xi_{p_2}(\hat{\tau}^{(j)}) - \text{LSL}\right)}{\left(\left(\xi_{p_3}(\hat{\tau}) - \xi_{p_1}(\hat{\tau})\right)\big/2\right)}; \quad j = 1, 2, \dots, B. \tag{16}$$

where $\hat{\tau}^{(j)}$ is the jth bootstrap estimator of τ.

Hence, the sample average and standard deviation are obtained as

$$\bar{\bar{C}}_{Npk}^{*} = \frac{1}{B}\sum_{j=1}^{B}\hat{C}_{Npk}^{*(j)} \quad \text{and} \quad S_{\hat{C}_{Npk}^{*}}^{*} = \sqrt{\frac{1}{B-1}\sum_{j=1}^{B}\left(\hat{C}_{Npk}^{*(j)} - \bar{C}_{Npk}^{*}\right)^{2}}.$$

Then, the 100 $(1 - \alpha)$% standard bootstrap (SB) confidence interval is

$$CI_{SB} = \left(\bar{\bar{C}}_{Npk}^{*} - Z_{1-\alpha/2}S_{\hat{C}_{Npk}^{*}}^{*}, \bar{\bar{C}}_{Npk}^{*} + Z_{1-\alpha/2}S_{\hat{C}_{Npk}^{*}}^{*}\right), \tag{17}$$

where $Z_{1-\alpha/2}$ is the $(1 - \alpha/2)$th quantile of the standard normal distribution.

Percentile bootstrap (PB) confidence interval

The 100 $(1 - \alpha)$% percentile bootstrap (PB) confidence interval is given by

$$CI_{PB} = \left(\hat{C}_{Npk(B(\alpha/2))}^{*}, \hat{C}_{Npk(B(1-\alpha/2))}^{*}\right), \tag{18}$$

where $\hat{C}_{Npk(B(\alpha/2))}^{*}$ and $\hat{C}_{Npk(B(1-\alpha/2))}^{*}$ be the $100(\alpha/2)$th and $100(1 - \alpha/2)$th empirical percentiles of $\hat{C}_{Npk}^{*}$, respectively.

Bias-corrected percentile bootstrap (BCPB) confidence interval

The complete bootstrap distribution obtained from a sample which could result in the higher or lower side of the expected value which is nothing but the bias, as the name indicates; the third method introduced to correct this potential bias is bias-corrected percentile bootstrap (BCPB) confidence interval. We use the ordered distribution of $\hat{C}_{Npk}^{*}$, to compute the probability $p_0 = \Pr\left[\hat{C}_{Npk}^{*} \leq \hat{C}_{Npk}\right]$, where $\hat{C}_{Npk}^{*}$ is the estimated value of $\hat{C}_{Npk}$. Then, the following are determined:

1. The biased-correction factor $Z_0 = \Phi^{-1}(p_0)$, where $\Phi(.)$ is the cumulative standard normal distribution function.
2. $P_L = \Phi(2Z_0 + Z_{\alpha/2})$ and $P_U = \Phi(2Z_0 + Z_{1-\alpha/2})$ are computed using Z_0 values.

Hence, 100 $(1 - \alpha)$% bias-corrected percentile bootstrap (BCPB) confidence interval for $\hat{C}_{Npk}$ is given by

$$CI_{BCPB} = \left(\hat{C}_{Npk(BP_L)}^{*}, \hat{C}_{Npk(BP_U)}^{*}\right), \tag{19}$$

where $\hat{C}_{Npk(r)}^{*}$ is the rth ordered value of the B bootstrap estimator of $\hat{C}_{Npk}$.

To determine the performance of the above three confidence intervals, we considered their estimated coverage probabilities and average widths. The probability that the true value of $\hat{C}_{Npk}$ is covered by the 100 $(1 - \alpha)$% bootstrap confidence interval is called "coverage probability," and the same is obtained for the methods discussed above. In addition, the average width of the bootstrap confidence interval is calculated based on 5000 different trials. The performance of the confidence intervals CI_{SB}, CI_{PB} and CI_{BCPB} based on their estimated coverage probabilities and average widths of the bootstrap confidence interval is studied through simulation.

Simulation study

The present section dealt with the results obtained through simulation study on the evaluation of three bootstrap confidence intervals of the process capability index given in Eq. (16) of TGLLD. With different parametric combinations $\sigma = 1$ and $\lambda = 4, 5, 6, 7$, we consider the sample size $n = 10, 15, 20, 25, 30$ and set the lower and upper specification limits as 1 and 29, respectively, to draw the simulation results. $B = 10000$ bootstrap samples of size n are generated from the original sample and repeated the exercise 5000 times. Using three methods, i.e., SB, PB and BCPB, the 95% confidence intervals were obtained. The difference between upper and lower confidence limits called the average width along with bias and MSE is calculated to compare the simulation results which are presented in Tables 1, 2, 3 and 4. The criteria set for comparison of results is the indices having lower average width, and higher coverage probabilities are to be considered.

It is noticed from the results given in Tables 1, 2, 3 and 4 that when the sample size grows the corresponding average width is falling down, indicating that moderately large sample throws better results. When we compared the average width values, BCPB method recorded lower values than SB and PB methods and followed the order BCPB < PB < SB. Average width of all the methods showed an upward trend when the shape parameter λ raises from 4 to 7. Similar pattern is observed when the other shape parameter θ turns up from 3.5 to 5. From the coverage probabilities recorded in Tables 1, 2, 3 and 4 for all three methods, a raising pattern is observed when λ increases from 4 to 7. SB method recorded higher estimated coverage probabilities which are more than

Table 1 Estimated coverage probabilities and average widths of 95% bootstrap confidence intervals of C_{Npk} for $\sigma = 1$ and $\lambda = 4$

n	θ	True C_{Npk}	Est C_{Npk}	Bias	MSE	Average widths			Coverage probabilities		
						SB	PB	BCPB	SB	PB	BCPB
10	3.5	0.9783	1.2504	−0.1590	0.1556	1.5998	1.5701	1.2448	0.9482	0.8326	0.9120
10	4	1.0167	1.2866	−0.1615	0.1569	1.6115	1.5824	1.2508	0.9496	0.8304	0.9104
10	4.5	1.0468	1.3151	−0.1630	0.1573	1.6189	1.5899	1.2546	0.9520	0.8308	0.9110
10	5	1.0712	1.3382	−0.1639	0.1571	1.6236	1.5947	1.2566	0.9536	0.8298	0.9116
15	3.5	0.9783	1.1952	−0.1024	0.0868	1.1473	1.1357	0.9831	0.9366	0.8656	0.9194
15	4	1.0167	1.2289	−0.1041	0.0878	1.1544	1.1429	0.9876	0.9356	0.8644	0.9214
15	4.5	1.0468	1.2553	−0.1053	0.0883	1.1586	1.1475	0.9900	0.9368	0.8640	0.9204
15	5	1.0712	1.2767	−0.1061	0.0885	1.1612	1.1503	0.9915	0.9362	0.8638	0.9212
20	3.5	0.9783	1.2329	−0.0717	0.0577	0.9365	0.9293	0.8372	0.9362	0.8880	0.9288
20	4	1.0167	1.2653	−0.0731	0.0584	0.9418	0.9347	0.8412	0.9358	0.8864	0.9282
20	4.5	1.0468	1.2906	−0.0741	0.0589	0.9448	0.9378	0.8434	0.9352	0.8852	0.9276
20	5	1.0712	1.3110	−0.0749	0.0591	0.9465	0.9396	0.8444	0.9348	0.8856	0.9266
25	3.5	0.9783	1.3073	−0.1872	0.0825	1.8363	1.8189	1.4176	0.9486	0.8334	0.9118
25	4	1.0167	1.3421	−0.0591	0.0429	0.8148	0.8091	0.7456	0.9454	0.8980	0.9394
25	4.5	1.0468	1.3690	−0.0598	0.0431	0.8171	0.8116	0.7472	0.9434	0.8992	0.9384
25	5	1.0712	1.3905	−0.0603	0.0432	0.8183	0.8129	0.7479	0.9426	0.8984	0.9384
30	3.5	0.9783	1.0222	−0.0883	0.0416	0.8103	0.8234	0.6879	0.9394	0.8900	0.9302
30	4	1.0468	1.0625	−0.0512	0.0370	0.7304	0.7255	0.6785	0.9316	0.8980	0.9266
30	4.5	1.0712	1.0942	−0.0516	0.0371	0.7313	0.7265	0.6791	0.9306	0.8984	0.9266
30	5	1.0912	1.1198	−0.0519	0.0371	0.7317	0.7269	0.6792	0.9292	0.8976	0.9250

Table 2 Estimated coverage probabilities and average widths of 95% bootstrap confidence intervals of C_{Npk} for $\sigma = 1$ and $\lambda = 5$

n	θ	True C_{Npk}	Est C_{Npk}	Bias	MSE	Average widths			Coverage probabilities		
						SB	PB	BCPB	SB	PB	BCPB
10	3.5	1.2260	1.5286	−0.1872	0.2125	1.8963	1.8589	1.4576	0.9586	0.8334	0.9118
10	4	1.2671	1.6899	−0.1906	0.2156	1.9126	1.8754	1.4645	0.9596	0.8322	0.9110
10	4.5	1.2994	1.6194	−0.1928	0.9114	1.9236	1.8860	1.4690	0.9600	0.8324	0.2163
10	5	1.3253	1.3138	−0.1939	0.2156	1.9309	1.8934	1.4716	0.9612	0.8312	0.9106
15	3.5	1.2260	1.1176	−0.1202	0.1174	1.3451	1.3308	1.1435	0.9436	0.8662	0.9208
15	4	1.2671	1.2416	−0.1225	0.1190	1.3546	1.3404	1.1485	0.9432	0.8650	0.9202
15	4.5	1.2994	1.0503	−0.1241	0.1200	1.3609	1.3469	1.1519	0.9418	0.8658	0.9220
15	5	1.3253	1.1173	−0.1253	0.1206	1.3652	1.3514	1.1539	0.9430	0.8640	0.9222
20	3.5	1.2260	2.2360	−0.0842	0.0774	1.0915	1.0830	0.9704	0.9402	0.8894	0.9294
20	4	1.2671	1.3674	−0.0860	0.0786	1.0984	1.0898	0.9749	0.9404	0.8868	0.9280
20	4.5	1.2994	1.9263	−0.0873	0.0793	1.1029	1.0944	0.9776	0.9402	0.8858	0.9274
20	5	1.3253	1.0339	−0.0883	0.0798	1.1058	1.0973	0.9791	0.9404	0.8864	0.9274
25	3.5	1.2260	0.9242	−0.0681	0.0567	0.9416	0.9349	0.8586	0.9490	0.8984	0.9412
25	4	1.2671	1.3628	−0.0693	0.0574	0.9470	0.9403	0.8622	0.9470	0.8990	0.9392
25	4.5	1.2994	1.5053	−0.0702	0.0578	0.9504	0.9437	0.8643	0.9466	0.8998	0.9394
25	5	1.3253	1.2306	−0.0708	0.0580	0.9524	0.9459	0.8654	0.9466	0.8988	0.9380
30	3.5	1.2260	1.0429	−0.0583	0.0486	0.8403	0.8345	0.7788	0.9394	0.9002	0.9306
30	4	1.2671	0.9676	−0.0593	0.0491	0.8448	0.8391	0.7821	0.9368	0.8992	0.9280
30	4.5	1.2994	1.3758	−0.0601	0.0495	0.8476	0.8419	0.7839	0.9354	0.8988	0.9266
30	5	1.3253	1.5513	−0.0606	0.0497	0.8493	0.8436	0.7848	0.9342	0.8990	0.9260

the confidence level (0.95) than BCPB and PB methods, and the pattern observed is PB < BCPB < SB. These probabilities get nearer to the confidence level (0.95) in BCPB method when the shape parameter λ turns up from 4 to 7. When the sample size increased to 30, bias and MSE recorded their lowest values at the parametric values $\lambda = 4$ and $\theta = 4$.

Table 3 Estimated coverage probabilities and average widths of 95% bootstrap confidence intervals of C_{Npk} for $\sigma = 1$ and $\lambda = 6$

n	θ	True C_{Npk}	Est C_{Npk}	Bias	MSE	Average widths			Coverage probabilities		
						SB	PB	BCPB	SB	PB	BCPB
10	3.5	1.4608	2.3146	− 0.2172	0.2822	2.2011	2.1551	1.6740	0.9642	0.8348	0.9110
10	4	1.5046	2.3543	− 0.2209	0.2849	2.2219	2.1756	1.6824	0.9644	0.8330	0.9118
10	4.5	1.5389	2.3837	− 0.2237	0.2872	2.2370	2.1899	1.6887	0.9664	0.8328	0.9106
10	5	1.5664	2.4054	− 0.2251	0.2867	2.2473	2.2002	1.6926	0.9674	0.8330	0.9142
15	3.5	1.4608	1.3953	− 0.1391	0.1542	1.5494	1.5317	1.3078	0.9462	0.8672	0.9222
15	4	1.5046	1.4521	− 0.1418	0.1567	1.5620	1.5444	1.3143	0.9458	0.8662	0.9216
15	4.5	1.5389	1.4972	− 0.1438	0.1583	1.5708	1.5536	1.3191	0.9458	0.8666	0.9230
15	5	1.5664	1.5337	− 0.1453	0.1594	1.5772	1.5599	1.3221	0.9464	0.8648	0.9224
20	3.5	1.4608	1.7916	− 0.0975	0.1012	1.2522	1.2417	1.1074	0.9438	0.8892	0.9294
20	4	1.5046	1.8281	− 0.0996	0.1029	1.2614	1.2508	1.1132	0.9444	0.8884	0.9282
20	4.5	1.5389	1.8566	− 0.1013	0.1040	1.2677	1.2572	1.1171	0.9438	0.8872	0.9280
20	5	1.5664	1.8796	− 0.1025	0.1049	1.2722	1.2617	1.1194	0.9442	0.8870	0.9282
25	3.5	1.4608	1.6327	− 0.0785	0.0739	1.0777	1.0696	0.9785	0.9502	0.8994	0.9412
25	4	1.5046	1.6916	− 0.0800	0.0749	1.0849	1.0768	0.9833	0.9496	0.9012	0.9394
25	4.5	1.5389	1.7374	− 0.0812	0.0756	1.0897	1.0816	0.9863	0.9494	0.9002	0.9392
25	5	1.5664	1.7739	− 0.0820	0.0761	1.0931	1.0852	0.9884	0.9494	0.8982	0.9380
30	3.5	1.4608	1.7294	− 0.0672	0.0632	0.9603	0.9535	0.8870	0.9404	0.9000	0.9296
30	4	1.5046	1.7707	− 0.0685	0.0640	0.9664	0.9594	0.8912	0.9384	0.8998	0.9278
30	4.5	1.5389	1.8031	− 0.0695	0.0646	0.9705	0.9635	0.8938	0.9372	0.9000	0.9264
30	5	1.5664	1.8294	− 0.0702	0.0650	0.9733	0.9662	0.8955	0.9368	0.8984	0.9258

Table 4 Estimated coverage probabilities and average widths of 95% bootstrap confidence intervals of C_{Npk} for $\sigma = 1$ and $\lambda = 7$

n	θ	True C_{Npk}	Est C_{Npk}	Bias	MSE	Average widths			Coverage probabilities		
						SB	PB	BCPB	SB	PB	BCPB
10	3.5	1.6878	2.6521	− 0.2481	0.3651	2.5120	2.4567	1.8947	0.9688	0.8352	0.9112
10	4	1.7344	2.6951	− 0.2529	0.3708	2.5386	2.4825	1.9060	0.9688	0.8344	0.9124
10	4.5	1.7709	2.7276	− 0.2560	0.3733	2.5575	2.5008	1.9140	0.9700	0.8342	0.9134
10	5	1.8001	2.7515	− 0.2575	0.3718	2.5707	2.5137	1.9192	0.9712	0.8342	0.9142
15	3.5	1.6878	1.6216	− 0.1585	0.1975	1.7588	1.7375	1.4761	0.9482	0.8682	0.9224
15	4	1.7344	1.6821	− 0.1617	0.2010	1.7748	1.7536	1.4847	0.9480	0.8668	0.9234
15	4.5	1.7709	1.7301	− 0.1641	0.2032	1.7863	1.7654	1.4910	0.9482	0.8666	0.9226
15	5	1.8001	1.7690	− 0.1659	0.2048	1.7949	1.7739	1.4953	0.9488	0.8658	0.9238
20	3.5	1.6878	1.7655	− 0.1111	0.1291	1.4174	1.4048	1.2483	0.9480	0.8896	0.9290
20	4	1.7344	1.8216	− 0.1137	0.1315	1.4291	1.4165	1.2558	0.9470	0.8888	0.9276
20	4.5	1.7709	1.8651	− 0.1156	0.1332	1.4375	1.4248	1.2609	0.9472	0.8882	0.9282
20	5	1.8001	1.8995	− 0.1171	0.1344	1.4437	1.4310	1.2643	0.9458	0.8858	0.9280
25	3.5	1.6878	1.8853	− 0.0893	0.0940	1.2178	1.2083	1.1020	0.9520	0.9008	0.9406
25	4	1.7344	1.9486	− 0.0911	0.0955	1.2271	1.2175	1.1083	0.9508	0.9020	0.9390
25	4.5	1.7709	1.9978	− 0.0925	0.0965	1.2336	1.2240	1.1125	0.9508	0.9004	0.9384
25	5	1.8001	2.0370	− 0.0935	0.0973	1.2384	1.2289	1.1156	0.9508	0.8990	0.9378
30	3.5	1.6878	1.9931	− 0.0765	0.0803	1.0841	1.0761	0.9983	0.9418	0.9008	0.9296
30	4	1.7344	2.0367	− 0.0781	0.0815	1.0920	1.0838	1.0039	0.9402	0.9002	0.9284
30	4.5	1.7709	2.0709	− 0.0792	0.0824	1.0975	1.0892	1.0076	0.9390	0.8996	0.9266
30	5	1.8001	2.0986	− 0.0801	0.0830	1.1015	1.0931	1.0103	0.9390	0.8980	0.9254

Illustrative example

In this section, we use a real data set to show that the type-II generalized log-logistic distribution can be a suitable model. Folks and Chhikara (1978) presented several sets of data to describe the Birnbaum–Saunders distribution. One of the data sets gives the runoff amounts at Jug Bridge, Maryland. For ready reference, this data set is reproduced as follows:

0.17, 0.23, 0.33, 0.39, 0.39, 0.40, 0.45, 0.52, 0.56, 0.59, 0.64, 0.66, 0.70, 0.76, 0.77, 0.78, 0.95, 0.97, 1.02, 1.12, 1.19, 1.24, 1.59, 1.74 and 2.92.

We show a rough indication of the goodness of fit for our model by plotting the superimposed for the data shows that the TGLLD is a good fit as shown in Fig. 1 and also goodness of fit is emphasized with QQ plot, as displayed in Fig. 1. The maximum likelihood estimates of the three-parameter TGLLD for the runoff amounts are $\hat{\sigma} = 0.7616$, $\hat{\lambda} = 2.6602$ and $\hat{\theta} = 1.772$; the Kolmogorov–Smirnov test found that the maximum distance between the data and the fitted TGLLD is 0.0657 with p value 0.9999. As it can be seen from the high p value that the data set considered is non-normal, TGLLD model is the best fit to this data set. Meanwhile, the maximum likelihood estimates of the two-parameter TGLLD for the runoff amounts are $\hat{\lambda} = 2.6602$ and $\hat{\theta} = 1.772$; the Kolmogorov–Smirnov test found that the maximum distance between the data and the fitted TGLLD is 0.2526 with p value 0.1891. Therefore, the two-parameter TGLLD also provides reasonable good fit for the runoff amounts.

The bootstrap confidence intervals and widths of C_{Npk} and C_{pk} are given in Table 5 for the given example. Numerical example shows that width of class intervals is considerably large in traditional C_{pk} method as compared to the bootstrap approach for C_{Npk}. Moreover, among the three bootstrap methods BCPB shows better performance than the other two methods; the simulation results also show the same.

Conclusions

In this article, we constructed bootstrap confidence intervals of process capability index using bootstrap method (proposed by Chen and Pearn, (1997)) by applying simulation technique, assuming that the underlying distribution is TGLLD. Bootstrap confidence intervals are constructed using three methods, i.e., SB, PB and BCPB. The performance of these methods is compared by deriving average

Table 5 Bootstrap confidence intervals and widths of new C_{Npk} and traditional C_{pk} for TGLLD

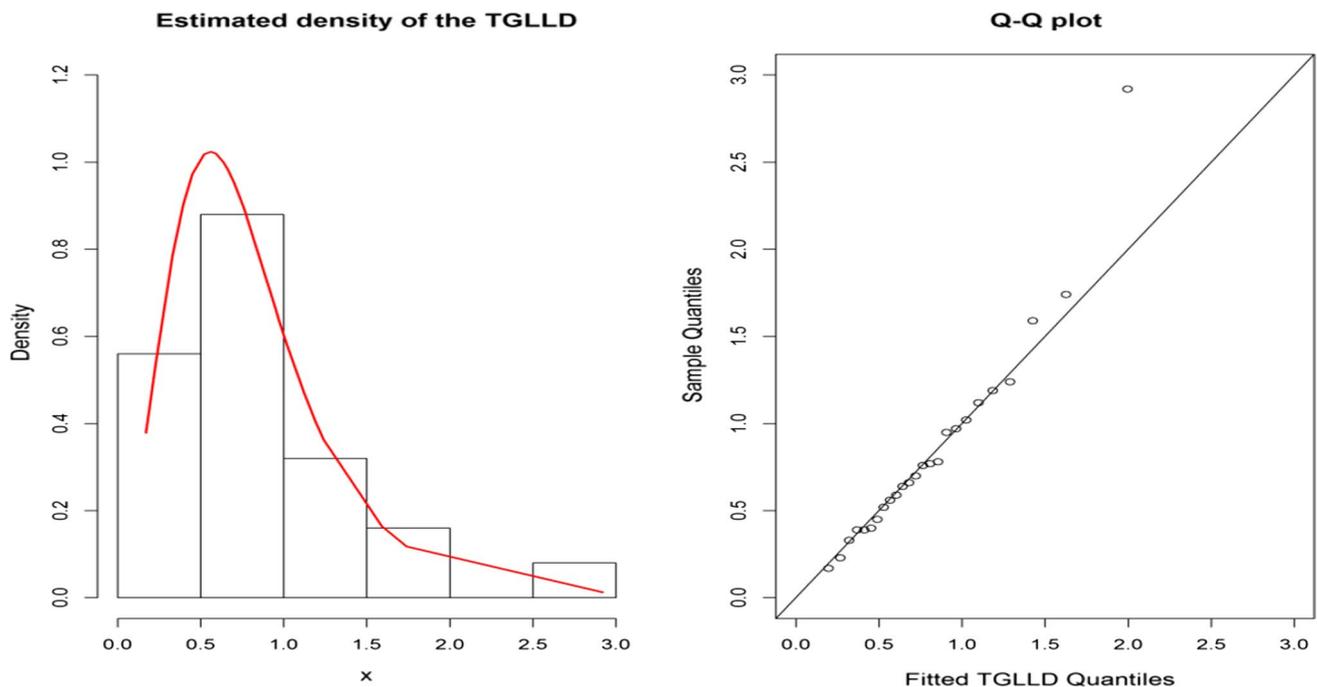| Methods	Bootstrap C_{Npk}		Traditional C_{pk}	
	Confidence intervals	Widths	Confidence intervals	Widths
SB	(0.0068, 0.1705)	0.1637	(0.2065, 0.5223)	0.3158
PB	(0.0379, 0.1992)	0.1613	(0.1248, 0.4405)	0.3156
BCPB	(0.0338, 0.1896)	0.1557	(0.1982, 0.4949)	0.2967

Fig. 1 Density plot and Q–Q plot of the fitted type-II generalized log-logistic distribution for the runoff amounts data

width, coverage probabilities, bias and MSE from simulation results. ML method of estimation is used to estimate the parameters under study. When both average width and coverage probabilities are considered as the performance criteria, BCPB method throws better results than the other two methods.

References

Aslam M, Wu CW, Azam M, Jun C-H (2013) Variable sampling inspection for resubmitted lots based on process capability index C_{pk} for normally distributed items. Appl Math Model 37(3):667–675

Chan LK, Cheng SW, Spring FA (1988) Anew measure of process capability C_{pm}. J Qual Technol 20(3):162–175

Chen KS, Pearn WL (1997) An application of non-normal process capability indices. Qual Reliab Eng Int 13(6):355–360

Chen KS, Huang ML, Hung YH (2008) Process capability analysis chart with the application of Cpm. Int J Prod Res 46(16):4483–4499

Clements JA (1989) Process capability calculations for non-normal distribution. Qual Prog 24(9):95–100

Efron B (1982) The jackknife, the bootstrap and other re-sampling plans, CBMS-NSF regional conference series in applied mathematics, 38. SIAM, Philadelphia

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall, Newyork

Folks JL, Chhikara R (1978) The inverse Gaussian distribution and its statistical application. J R Stat Soc Ser B Stat Method 20:263–289

Gunter BH (1989a) The use and abuse of C, parts 1–4. Qual Prog 24(1):72–73

Gunter BH (1989b) The use and abuse of C, parts 1–4. Qual Prog 24(3):108–109

Gunter BH (1989c) The use and abuse of C, parts 1–4. Qual Prog 24(5):79–80

Gunter BH (1989d) The use and abuse of C, parts 1–4. Qual Prog 24(7):86–87

Juran JM (1974) Juran's quality control handbook. McGraw Hill, New York

Kane VE (1986) Process capability indices. J Qual Technol 18(51):41052

Kantam RRL, Rosaiah K, Subba Rao R (2010) Estimation of process capability index for half logistic distribution. Int Trans Math Sci Comput 3(1):61–66

Kashif M, Aslam M, Rao GS, AL-Marshadi AH, Jun CH (2017) Bootstrap confidence intervals of the modified process capability index for Weibull distribution. Arab J Sci Eng 42:4565–4573

Pearn WL, Chen KS (1995) Estimating process capability indices for non-normal Pearson populations. Qual Reliab Eng Int 11(5):386–388

Peng C (2010a) Parametric lower confidence limits of quantile-based process capability indices. Qual Technol Quant Manag 7(3):199–214

Peng C (2010b) Estimating and testing quantile-based process capability indices for processes with skewed distributions. J Data Sci 8(2):253–268

Perakis M (2010) Estimation of differences between process capability indices c_{pm} or c_{pmk} for two processes. J Stat Comput Simul 80(3):315–334

Rao GS, Kantam RRL, Rosaiah K, Prasad SVSVSV (2012a) Reliability test plans for type-II generalized log-logistic distribution. JRSS 5(1):55–64

Rao GS, Kantam RRL, Rosaiah K, Prasad SVSVSV (2012b) An economic reliability test plan for generalized log-logistic distribution. Int J Eng Appl Sci 3:61–68

Rao GS, Aslam M, Kantam RRL (2015) Bootstrap confidence intervals of c_{Npk} for inverse Rayleigh and log-logistic distributions. J Stat Comput Simul 86(5):862–873

Rosaiah K, Kantam RRL, Santosh Kumar Ch (2006) Reliability of test plans for exponentiated log-logistic distribution. Econ Qual Control 21(2):165–175

Rosaiah K, Kantam RRL, Santosh Kumar Ch (2007) Exponentiated log- logistic distribution—an economic reliability test plan. Pak J Stat 23(2):147–156

Rosaiah K, Kantam RRL, Prasad SVSVSV, Pratapa Reddy J (2008) Reliability estimation in type-II generalized log-logistic distribution. Int J Agric Stat Sci 4(2):283–292

Vännman KA (1995) Unified approach to process capability indices. Stat Sin 5(2):805–820

Wararit P, Somchit W (2012) Bootstrap confidence intervals of the difference between two process capability indices for half logistic distribution. Pak J Stat Oper Res 8(4):879–894

Wood M (2005) Bootstrapped confidence intervals as an approach to statistical inference. Res Methods 8(4):454–470

Wu SF, Liang MC (2010) A note on the asymptotic distribution of the process capability index C_{pmk}. J Stat Comput Simul 80(2):227–235