정규분포 공정 가정하에서의 공정능력지수 C_{pmk}에 관한 효율적인 신뢰한계*

조중재*† · 박병선** · 박효일***

*충북대학교 자연과학대학 정보통계학과

**한국표준과학연구원 측정품질그룹

***청주대학교 이공대학 통계학과

Better Confidence Limits for Process Capability Index C_{pmk} under the assumption of Normal Process

Joong-Jae Cho* · Byoung-Sun Park** · Hyo-il Park***

*Department of Statistics, Chungbuk National University

**Measurement & Quality Service Group, Korea Research Institute of Standards and Science

***Department of Statistics, Cheongju University

Key Words : Process capability index, Asymptotic distribution, Bootstrap consistency, Monte-carlo simulation, Statistical inference, Bootstrap confidence limits

Abstract

Process capability index is used to determine whether a production process is capable of producing items within a specified tolerance. The index C_{pmk} is the third generation process capability index. This index is more powerful than two useful indices C_p and C_{pk}. Whether a process distribution is clearly normal or nonnormal, there may be some questions as to which any process index is valid or should even be calculated. As far as we know, yet there is no result for statistical inference with process capability index C_{pmk}. However, asymptotic method and bootstrap could be studied for good statistical inference. In this paper, we propose various bootstrap confidence limits for our process capability index C_{pmk}. First, we derive bootstrap asymptotic distribution of plug-in estimator $\hat{C}_{pmk}$ of our capability index C_{pmk}. And then we construct various bootstrap confidence limits of our capability index C_{pmk} for more useful process capability analysis.

† 교신저자 jjcho@chungbuk.ac.kr

* This work was supported by Chungbuk National University Grant in 2004.

1. Introduction

Process capability indices, whose purpose is to provide a numerical measure on whether a production process is capable of producing items satisfying the quality requirements preset by the designer, have received substantial attention in the quality control and statistical literature. The three most widely used capability indices are as follows.

$$C_p = \frac{USL-LSL}{6\sigma}$$

$$C_{pk} = \min\left\{\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right\}$$

$$C_{pm} = \frac{USL-LSL}{6\sqrt{E(X-T)^2}}$$

where USL is the upper specification limit and LSL is the lower specification limit. Also, μ is the process mean, σ is the process statndard deviation, and T is the target value.

While the index C_p reflects only the magnitude of the process variation, the index C_{pk} takes into account the process variation as well as the location of the process mean relative to the specification limits. Also, to obtain more sensitive capability index than C_{pk} and C_{pm}, Pearn et al.(1992) introduced the third process capability index C_{pmk} as follows.

$$C_{pmk} = \frac{\min(USL-\mu, \mu-LSL)}{3\sqrt{E(X-T)^2}} = \frac{d-|\mu-M|}{3\sqrt{\sigma^2+(\mu-T)^2}}$$

where two constants d and M are defined by $d=(USL-LSL)/2$ and $M=(USL+LSL)/2$, respectively.

In general, the calculation of various lower confidence limits assumes a normally distributed process, and, as Gunter(1989) has noted, many real world processes are not normally distributed and this departure from normality may be hard to detect. This could potentially affect both the estimates of the indices and the lower confidence limits based on these estimates. Efron(1979) introduced and developed the nonparametric, but computer intensive, estimation method called bootstrap. In particular, Efron and Tibshirani(1986) further develop three types of bootstrap confidence intervals: the standard bootstrap confidence interval (SB), the percentile bootstrap confidence interval (PB), and the biased-corrected percentile bootstrap confidence interval (BCPB). Franklin and Wasserman(1991) presented an inital study of the properties of these three bootstrap confidence intervals for C_{pk}. Also, Franklin and Wasserman(1992) have studied bootstrap lower confidence limits for capability indices. However, there is no result on statistical inference for the index C_{pmk} because of computing complexity.

In this paper, we study bootstrap confidence limits for our process capability index C_{pmk}. First, we have derived bootstrap asymptotic distribution for our capability index C_{pmk}. Having provided the consistency of our bootstrap for process capability index C_{pmk}, we construct six bootstrap confidence intervals including three types of bootstrap confidence intervals SB, PB, and BCPB abovementioned. These results will play an important role to study their performances for our process index C_{pmk} under nonnormal distributions.

2. Bootstrapping our process capability index

2.1 Bootstrap algorithm

In this section, we introduce the bootstrap algorithm for deriving asymptotic distributions and confidence limits with the bootstrap. Some process capability indices(PCIs) are used to determine whether a production process is capable of producing items within a specified tolerance. They are considered as a practical tool by several advocates of statistical process control in industry.

Suppose that a set of the independent random variables $X_1, X_2, \cdots, X_n$ has a common distribution $F(\cdot)$ with process mean μ and process variance σ^2.

We obtain the natural estimators of C_{pmk} with the plug-in principle as follows.

$$\hat{C}_{pmk} = \frac{\min(USL-\overline{X},\ \overline{X}-LSL)}{3\sqrt{S^2+(\overline{X}-T)^2}},$$
$$= \frac{d-|\overline{X}-M|}{3\sqrt{S^2+(\overline{X}-T)^2}}$$

where sample mean $\overline{X}$ and sample variance S^2 implies $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ and $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ respectively.

Putting mass $1/n$ at each of the points $X_1, X_2, \cdots, X_n$, we get the bootstrap sample of size m, $X_1^*, X_2^*, \cdots, X_m^*$.

The bootstrap method is to approximate the distribution of $t(X_1, X_2, \cdots, X_n; F)$ under F by that of $t(X_1^*, X_2^*, \cdots, X_m^*; F_n)$. A formal description of the bootstrap algorithm goes as follows.

- Step 1 : Given $\chi_n = (X_1, X_2, \cdots, X_n)$, the bootstrap sample $X_1^*, X_2^*, \cdots, X_m^*$ can be obtained with replacement, which is conditionally independent with common distribution F_n.
- Step 2 : From the bootstrap sample $X^*_1, X^*_2 \cdots, X^*_m$, compute the bootstrap sample mean $\overline{X^*}$ and bootstrap sample variance S^{*2} as follows.

$$\overline{X^*} = \frac{1}{m}\sum_{i=1}^{m}X_i^*,$$
$$S^{*2} = \frac{1}{m-1}\sum_{i=1}^{m}(X_i^*-\overline{X^*})^2$$

- Step 3 : Compute the bootstrap plug-in estimator of C_{pmk}, as follows;

$$\hat{C}_{pmk}^{\ *} = \frac{\min(USL-\overline{X}^*,\ \overline{X}^*-LSL)}{3\sqrt{S^{*2}+(\overline{X}^*-T)^2}} = \frac{d-|\overline{X}^*-M|}{3\sqrt{S^{*2}+(\overline{X}^*-T)^2}}$$

2.2 Bootstrap asymptotic distribution

Now we study the asymptotic properties which are needed to construct various bootstrap confidence intervals for the process capability index C_{pmk}.

First, we introduce the asymptotic result for our capability index C_{pmk}, as follows.

Theorem 1 Assume that $\mu_4 = E(X-\mu)^4$ exists. Along almost all sample sequences given $X_n = (X_1, X_2, \cdots, X_n)$, as n tends to ∞, the following result holds:

$$\sqrt{n}(\hat{C}_{pmk} - C_{pmk})$$

$$\xrightarrow{d} \begin{cases} N(0,\ \sigma^2_{pmk}), & \mu < M \\ -\frac{|Y|}{3\tau} - \frac{dZ}{6\tau^3} + \frac{dY(T-\mu)}{3\tau^3}, & \mu = M \\ N(0, \sigma^{2'}_{pmk}), & \mu > M \end{cases}$$

where

$$\tau = \sqrt{\sigma^2+(\mu-T)^2},$$

$$\sigma^2_{pmk} = \frac{1}{9\tau^6}\Big[\ \tau^2+(T-\mu)(d-M+\mu)^2\sigma^2 - \mu_3(d-M+\mu)\times\{\tau^2+(d-M+\mu)(T-\mu)\} + \frac{1}{4}(\mu_4-\sigma^4)(d-M+\mu)^2\Big],$$

$$(Y,Z) \sim BN((0,0),\Sigma)$$

$$\sigma^{2'}_{pmk} = \frac{1}{9\tau^6}\Big[\ \tau^2+(\mu-T)(d-\mu+M)^2\sigma^2 - \mu_3(d-\mu+M)\times\tau^2+(d-\mu+M)(\mu-T) + \frac{1}{4}(\mu_4-\sigma^{4})(d-\mu+M)^2\Big].$$

Proof See Chen and Hsu(1995).

Lemma 1 Along almost all sample sequences given $\chi_n = (X_1, X_2, \cdots, X_n)$, as m and n tend to ∞, we obtain:

$$\sqrt{m}(\overline{X^*}-\overline{X},\ S^{*2}-S^2)|\ \chi_n \xrightarrow{d} BN\left((0,0), \begin{pmatrix}\sigma^2 & \mu_3 \\ \mu_3 & \mu_4-\sigma^4\end{pmatrix}\right)$$

Proof Let F_n be the empirical distribution of $\binom{X_1}{X_1^2}, \binom{X_2}{X_2^2}, \cdots, \binom{X_n}{X_n^2}$. Given $\binom{X_1}{X_1^2}, \binom{X_2}{X_2^2}, \cdots, \binom{X_n}{X_n^2}$, let $\binom{X_1^*}{X_1^{*2}}, \binom{X_2^*}{X_2^{*2}}, \cdots, \binom{X_m^*}{X_m^{*2}}$ be conditionally independent, with common distribution F_n. With Bickel and Freedman(1981: Theorem1 and Theorem2) and Mallows(1972), we obtain the following limiting distribution. As n and m tend to ∞:

$$\sqrt{m}\left(\overline{X}^*-\overline{X}, \frac{1}{m}\sum_{i=1}^{m}X_i^{*2}-\frac{1}{n}\sum_{i=1}^{n}X_i^2\right)\Big|\ \chi_n \xrightarrow{d} BN\left((0,0), \begin{pmatrix}\sigma^2 & \mu_3+2\mu\sigma^2 \\ \mu_3+2\mu\sigma^2 & \mu_4+4\mu\mu_3+4\mu^2\sigma^2-\sigma^4\end{pmatrix}\right)$$

Hence, we obtain Lemma 1 by the above result with simple calculations.

Lemma 2 Assume that function $g(u, v)$ is differentiable. Along almost all sample sequences given $\chi_n = (X_1, X_2, \cdots, X_n)$, as m and n tend to ∞:

$$\sqrt{m}(g(\overline{X^*}, S^{*2}) - g(\overline{X}, S^2))|\chi_n \xrightarrow{d} N(0, D'\Sigma D)$$

where

$$D' = \left(\frac{\partial g(u,v)}{\partial u}\Big|_{\mu,\sigma^2},\ \frac{\partial g(u,v)}{\partial v}\Big|_{\mu,\sigma^2}\right) \neq (0,0).$$

Proof The Lemma2 follows from Lemma1 and the TheoremA(p.122) of Serfling(1980).

Also, the consistency of our bootstrap for statistical inference of our process capability index $\hat{C}_{pmk}$ is easily provided as follows.

Theorem 2 Assume that $\mu_4 = E(X-\mu)^4$ exists. Along almost all sample sequences given $X_n = (X_1, X_2, \cdots, X_n)$, with our bootstrap algorithm as n and m tend to ∞, we obtain as follows;

$$\sqrt{n}(\hat{C}^*_{pmk} - \hat{C}_{pmk})|\chi_n \xrightarrow{d} \begin{cases} N(0, \sigma^2_{pmk}), & \mu < M \\ -\frac{|Y|}{3\tau} - \frac{dZ}{6\tau^3} + \frac{dY(T-\mu)}{3\tau^3}, & \mu = M \\ N(0, \sigma^{2'}_{pmk}), & \mu > M \end{cases}$$

where

$$\tau = \sqrt{\sigma^2 + (\mu - T)^2},$$

$$\sigma^2_{pmk} = \frac{1}{9\tau^6}\Big[\Big\{\tau^2 + (T-\mu)(d-M+\mu)\Big\}^2\sigma^2 - \mu_3(d-M+\mu)\times\{\tau^2 + (d-M+\mu)(T-\mu)\} + \frac{1}{4}(\mu_4 - \sigma^4)(d-M+\mu)^2\Big],$$

$$(Y, Z) \sim BN((0,0), \Sigma)$$

$$\sigma^{2'}_{pmk} = \frac{1}{9\tau^6}\Big[\Big\{\tau^2 + (\mu - T)(d-\mu+M)\Big\}^2\sigma^2 - \mu_3(d-\mu+M)\times\{\tau^2 + (d-\mu+M)(\mu - T)\} + \frac{1}{4}(\mu_4 - \sigma^4)(d-\mu+M)^2\Big]$$

Proof The proof is obtained by applying Lemma1 and Lemma2 in case of $\mu < M$ or $\mu > M$. Also, the result of the case $\mu = M$ is derived by the following calculations with some limit theorems containing the Slutsky's theorem. We consider the case2 as follows:

$$\sqrt{m}(\hat{C}_{pmk^*} - \hat{C}_{pmk})|\chi_n$$
$$= \sqrt{m}\left(\frac{d - |\overline{X}^* - \mu|}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}} - \frac{d - |\overline{X} - \mu|}{3\sqrt{S^2 + (\overline{X} - T)^2}}\right)\Bigg|\chi_n$$
$$= \sqrt{m}\left(\frac{d}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}} - \frac{d}{3\sqrt{S^2 + (\overline{X} - T)^2}}\right)\Bigg|\chi_n$$
$$-\sqrt{m}\left(\frac{|\overline{X}^* - \mu|}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}} - \frac{|\overline{X} - \mu|}{3\sqrt{S^2 + (\overline{X} - T)^2}}\right)\Bigg|\chi_n$$

The first term is calculated as follows:

$$\sqrt{m}\left(\frac{d}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}} - \frac{d}{3\sqrt{S^2 + (\overline{X} - T)^2}}\right)\Bigg|\chi_n$$
$$= \frac{-\sqrt{m}d\left(\sqrt{S^{*2} + (\overline{X}^* - T)^2} - \sqrt{S^2 + (\overline{X} - T)^2}\right)}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}\sqrt{S^2 + (\overline{X} - T)^2}}\Bigg|\chi_n$$
$$= \frac{-\sqrt{m}d[S^{*2} - S^2 + (\overline{X}^* - T + \overline{X} - T)(\overline{X}^* - \overline{X})]}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}\sqrt{S^2 + (\overline{X} - T)^2}} \times \frac{1}{(\sqrt{S^{*2} + (\overline{X}^* - T)^2} + \sqrt{S^2 + (\overline{X} - T)^2})}\Bigg|\chi_n$$
$$\xrightarrow{d} (Y, Z) \text{ as } n \to \infty, m \to \infty$$

where $(Y, Z) \sim BN((0,0), \Sigma), \tau^2 = \sigma^2 + (\mu - T)^2$.
Also, the second term can be calculated as follows:

$$-\sqrt{m}\left(\frac{|\overline{X}^* - \mu|}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}} - \frac{|\overline{X} - \mu|}{3\sqrt{S^2 + (\overline{X} - T)^2}}\right)\Bigg|\chi_n$$
$$= \frac{-\sqrt{m}\left(|\overline{X}^* - \mu|\sqrt{S^2 + (\overline{X} - T)^2} - |\overline{X} - \mu|\sqrt{S^{*2} + (\overline{X}^* - T)^2}\right)}{3\sqrt{S^{*2} + (\overline{X}^* - T)^2}\sqrt{S^2 + (\overline{X} - T)^2}}\Bigg|\chi_n$$
$$\xrightarrow{d} -\frac{|Y|}{3\tau}$$

as $m \to \infty$ and $n \to \infty$ by rationalizing the numerator and applying some limit theorems to it. Above two results imply Theorem2 for the case $\mu = M$ immediately. This completes the proof.

Of course, these limiting distributions are identical with those of **Theorem 1.** This result is called the consistency of the bootstrap.

We allow the resample size m to differ from the number n of data points, to estimate the distribution of the bootstrap pivotal quantity, say,

$$Q^*_m = \sqrt{m}(\hat{C}^*_{pmk} - \hat{C}_{pmk}) / S^*_{pmk}\big|\chi_n$$

where S^{*2}_{pmk} is the bootstrap version of the plug-in estimator S^2_{pmk} of the variance σ^2_{pmk}.

In the resampling, the n data points $X_1, X_2, \cdots, X_n$ are treated as a population, with distribution function F_n and mean

$\overline{X}$ and $\overline{X^*}$ is considered as an bootstrap estimator of $\overline{X}$. First, take $m=n$. The idea is that the behavior of the bootstrap pivotal quantity Q_n^* mimics that of Q_n. Thus, the distribution of Q_n^* could be computed from the data and used to approximate the unknown sampling distribution of Q_n or even more directly, the bootstrap distribution of

$$\sqrt{n}(\hat{C}^*_{pmk} - \hat{C}_{pmk})|\ \chi_n$$

could be used to approximate the sampling distribution of $\sqrt{n}(\hat{C}_{pmk} - C_{pmk})$. Either approach would be lead to confidence intervals for C_{pmk} and would be useful if the bootstrap approximation were valid.

3. Bootstrap confidence limits

In this section, we construct six bootstrap confidence limits for our index C_{pmk} including Studentized Bootstrap (STUD), Hybrid Bootstrap(HYB), and Accelerated Bias-Corrected Bootstrap (ABC), which were theoretically studied by Hall(1988).

Construction of a two-sided $(1-2\alpha)100\%$ confidence interval will be described, and a lower $(1-\alpha)100\%$ confidence limit can be obtained by using only the lower limit. The bootstrap confidence intervals of C_{pmk} are easily obtained as follows. For each simulation study, a sample of size n was drawn and for each of size n, B bootstrap resamples(in this paper, we use B=1,000) were drawn from that single sample with our bootstrap algorithm. This single simulation was then replicated N times(in this paper, we use N=1,000).

3.1 Standard Bootstrap(SB) method

From the $B=1000$ bootstrap estimates, $\hat{C}^*_{pmk}(i)$, we calculate sample mean and sample standard deviation as belows.

$$\hat{C}^*_{pmk}(\cdot) = \frac{1}{B}\sum_{i=1}^{B}\hat{C}^*_{pmk}(i),$$

$$S^*_{C_{pmk}} = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}(\hat{C}^*_{pmk}(i) - \hat{C}^*_{pmk}(\cdot))^2}$$

In fact, since the distribution of $\hat{C}_{pmk}$ is approximately normal, we obtain the $(1-2\alpha)100\%$ SB confidence interval for C_{pmk}

$$(\hat{C}_{pmk} - z_{1-\alpha}S^*_{C_{pmk}}\quad \hat{C}_{pmk} + z_{1-\alpha}S^*_{C_{pmk}})$$

where $z_{1-\alpha}$ is the $(1-\alpha)$ quantile of the standard normal distribution.

3.2 Percentile Bootstrap(PB) method

The PB method is used in more than half of the papers on bootstrap confidence intervals. But Hall(1988) pointed out a criticism of the PB method. From the ordered collection of

$$\hat{C}^*_{pmk}(i),\ (\hat{C}^*_{pmk}(1) \le \cdots \le \hat{C}^*_{pmk}(B))$$

,

we obtain the $(1-2\alpha)100\%$ PB confidence interval for C_{pmk}

$$\left(\widehat{C}^*_{pmk}([\alpha B]), \quad \widehat{C}^*_{pmk}([(1-\alpha)B]) \right),$$

where [ ] implies the gauss bracket; that is, $[x]$ is the largest integer less than or equal to the real number x.

3.3 Biased-Corrected Percentile Bootstrap(BCPB) method

Our bootstrap estimator $\widehat{C}^*_{pmk}$ may be biased. The BCPB method has been developed to correct for this potential bias.

First, using the ordered distribution of $\widehat{C}^*_{pmk}$, calculate the probability

$$p_0 = P(\widehat{C}^*_{pmk} \le \widehat{C}_{pmk} | \chi_n).$$

Second, calculate the quantile point z_0 and probabilities, P_L and P_U such that the following equations are satisfied.

$$\begin{aligned} z_0 &= \phi^{-1}(p_0), \\ P_L &= \Phi(2z_0 - z_{1-\alpha}), \\ P_U &= \Phi(2z_0 + z_{1-\alpha}) \end{aligned}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Finally, the $(1-2\alpha)100\%$ BCPB confidence interval of the index C_{pmk} is

$$\left(\widehat{C}^*_{pmk}([P_L B]), \quad \widehat{C}^*_{pmk}([P_U B]) \right).$$

3.4 STUDentized bootstrap (STUD) method

Hall(1988) mentioned that the STUD method led to intervals which tended to be conservative in the sense that they had greater length and greater coverage than other competitors. So the STUD method does a better job than several other methods, provided that the variance estimate is chosen well. But this generalization can be failed in the case of distributions with exceptionally large positive kurtosis.

If process variance σ^2 is unknown, we define the reasonable critical points, $\hat{y}'_{pmk,\alpha}$, $\hat{y}_{pmk,\alpha}$ and $\hat{y}''_{pmk,\alpha}$ as follows.

$$P\left\{ \frac{\sqrt{m}(\widehat{C}_{pmk^*} - \widehat{C}_{pmk})}{\hat{\sigma}^{*\prime}_{pmk}} \le \hat{y}'_{pmk,\alpha} \middle| \chi_n \right\} = \alpha,$$

$$P\left\{ \frac{\sqrt{m}(\widehat{C}_{pmk^*} - \widehat{C}_{pmk})}{\hat{\sigma}^{*}_{pmk}} \le \hat{y}_{pmk,\alpha} \middle| \chi_n \right\} = \alpha,$$

$$P\left\{ \frac{\sqrt{m}(\widehat{C}_{pmk^*} - \widehat{C}_{pmk})}{\hat{\sigma}^{*\prime\prime}_{pmk}} \le \hat{y}''_{pmk,\alpha} \middle| \chi_n \right\} = \alpha.$$

Here $\hat{\sigma}^{*\prime}_{pmk}$, $\hat{\sigma}^{*}_{pmk}$ and $\hat{\sigma}^{*\prime\prime}_{pmk}$ are the bootstrap version of the standard deviation $\hat{\sigma}'_{pmk}$, $\hat{\sigma}_{pmk}$ and $\hat{\sigma}''_{pmk}$.

Then we construct the $(1-2\alpha)100\%$ STUD confidence interval represented by

ⓐ $LSL < \mu < M$,

$$\left(\widehat{C}_{pmk} - \frac{\hat{\sigma}_{pmk}}{\sqrt{n}} \hat{y}_{pmk,1-\alpha}, \quad \widehat{C}_{pmk} - \frac{\hat{\sigma}_{pmk}}{\sqrt{n}} \hat{y}_{pmk,\alpha} \right)$$

ⓑ $\mu = M$,

$$\left(\widehat{C}_{pmk} - \frac{\hat{\sigma}_{pmk}}{\sqrt{n}} \hat{y}_{pmk,1-\alpha}, \quad \widehat{C}_{pmk} - \frac{\hat{\sigma}_{pmk}}{\sqrt{n}} \hat{y}_{pmk,\alpha} \right)$$

ⓒ $M < \mu < USL$,

$$\left(\widehat{C}_{pmk} - \frac{\hat{\sigma}''_{pmk}}{\sqrt{n}} \hat{y}''_{pmk,1-\alpha}, \quad \widehat{C}_{pmk} - \frac{\hat{\sigma}''_{pmk}}{\sqrt{n}} \hat{y}''_{pmk,\alpha} \right)$$

where

$$\hat{\tau} = \sqrt{S^2 + (\overline{X} - T)^2},$$

$$\hat{\sigma}_{pmk}' = \Big[\frac{1}{9\hat{\tau}^6}\Big[\Big\{\hat{\tau}^2+(T-\overline{X})(d-M+\overline{X})\Big\}^2 S^2 - \hat{\mu}_3(d-M+\overline{X})\times\{\hat{\tau}^2+(d-M+\overline{X})(T-\overline{X})\} + \frac{1}{4}(\hat{\mu}_4-S^4)(d-M+\overline{X})^2\Big]\Big]^{1/2},$$

$$\hat{\sigma}_{pmk} = \Big[\frac{(\hat{\mu}_4-S^4)d^2}{36S^6}\Big]^{1/2},$$

$$\hat{\sigma}_{pmk}'' = \Big[\frac{1}{9\hat{\tau}^6}\Big[\Big\{\hat{\tau}^2+(\overline{X}-T)(d-\overline{X}+M)\Big\}^2 S^2 - \hat{\mu}_3(d-\overline{X}+M)\times\{\hat{\tau}^2+(d-\overline{X}+M)(\overline{X}-T)\} + \frac{1}{4}(\hat{\mu}_4-S^4)(d-\overline{X}+M)^2\Big]\Big]^{1/2},$$

3.5 HYBrid bootstrap(HYB) method

The HYB method is used in almost all of the studies not using the PB method. Some users are even unaware that there is a difference between the HYB and PB method. Equal-tailed intervals based on the HYB method and the PB method always have exactly the same length, but usually have different centers.

When process variance σ^2 is unknown, we can use the critical point, $\hat{x}_\alpha$ of the equation (4). The $(1-2\alpha)100\%$ HYB confidence interval is

ⓐ $LSL<\mu<M$,

$$\Big(\hat{C}_{pmk}-\frac{\hat{\sigma}'_{pmk}}{\sqrt{n}}\hat{x}'_{pmk,1-\alpha},\ \hat{C}_{pmk}-\frac{\hat{\sigma}'_{pmk}}{\sqrt{n}}\hat{x}'_{pmk,\alpha}\Big)$$

ⓑ $\mu=M$,

$$\Big(\hat{C}_{pmk}-\frac{\hat{\sigma}_{pmk}}{\sqrt{n}}\hat{x}_{pmk,1-\alpha},\ \hat{C}_{pmk}-\frac{\hat{\sigma}_{pmk}'}{\sqrt{n}}\hat{x}_{pmk,\alpha}\Big)$$

ⓒ $M<\mu<USL$,

$$\Big(\hat{C}_{pmk}-\frac{\hat{\sigma}''_{pmk}}{\sqrt{n}}\hat{x}''_{pmk,1-\alpha},\ \hat{C}_{pmk}-\frac{\hat{\sigma}''_{pmk}}{\sqrt{n}}\hat{x}''_{pmk,\alpha}\Big)$$

where $\hat{x}'_{pmk,\alpha}$, $\hat{x}_{pmk,\alpha}$ and $\hat{x}''_{pmk,\alpha}$ are satisfied the following equations, repectively.

$$P\Big\{\frac{\sqrt{m}(\hat{C}_{pmk^*}-\hat{C}_{pmk})}{\hat{\sigma}'_{pmk}} \le \hat{x}'_{pmk,\alpha}\Big|\chi_n\Big\} = \alpha,$$

$$P\Big\{\frac{\sqrt{m}(\hat{C}_{pmk^*}-\hat{C}_{pmk})}{\hat{\sigma}_{pmk}} \le \hat{x}_{pmk,\alpha}\Big|\chi_n\Big\} = \alpha,$$

$$P\Big\{\frac{\sqrt{m}(\hat{C}_{pmk^*}-\hat{C}_{pmk})}{\hat{\sigma}''_{pmk}} \le \hat{x}''_{pmk,\alpha}\Big|\chi_n\Big\} = \alpha.$$

3.6 Accelerated Bias-Corrected bootstrap(ABC) method

It may occur that the bootstrap distributions obtained by using only a sample of the complete bootstrap distribution may be shifted higher or lower than would be expected. Applied statisticians make frequent use of devices like transformations, bias corrections, and even acceleration adjustments, to improve the performance of the standard intervals. The ABC method enjoys useful properties of invariance under transformations, properties not shared by the STUD method, although the STUD method does a better job than the ABC method, provided the variance estimate is chosen correctly.

The acceleration constant, a , always measures the rate of change of standard error on a normalized scale. Consider $d=2$ and the function $g(\mu)=g(\mu_1,\mu_2)$ for process capability index C_{pmk}. The acceleration constant a for process indices defined by Hall(1998) is as follows.

$$a=\frac{1}{6\sqrt{n}\ \hat{\sigma}^3_{pmk}}\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}a_i a_j a_k \mu_{ijk}\right]$$

where $a_i=\partial g/\partial \mu_i,\ i=1,2$.

First, calculate the bootstrap estimator, $\hat{a}$, of the acceleration constant a for our capability index C_{pmk} as follows :

$$\hat{a}=\frac{1}{6\sqrt{n}\ \hat{\sigma}^3_{pmk}}\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\hat{a}_i\ \hat{a}_j\ \hat{a}_k\ \hat{\mu}_{ijk}\right].$$

We must keep in mind all estimators for acceleration contant a consist of bootstrap samples. This estimate, $\hat{a}$,approximately coincides with

$$\hat{a}\approx\frac{1}{6}\,Skew_{c=\hat{c}}\ (\dot{l}_c)$$

where $\dot{l}_c$ is the score function, which is given by Efron(1987).

This sounds different to compute, but it is in fact easier to get a good estimate of $\hat{a}$ than of z_0 .

Second, esitmate $\hat{\beta}_{aL}$ and $\hat{\beta}_{aU}$ which are

$$\hat{\beta}_{aL}=\Phi(z_{\alpha}+2z_0+\hat{a}z^2_{\alpha}),$$

$$\hat{\beta}_{aU}=\Phi(z_{1-\alpha}+2z_0+\hat{a}\,z^2_{1-\alpha}).$$

The $(1-2\alpha)100\%$ ABC confidence interval is

ⓐ $LSL<\mu<M,$

$$\left(\hat{C}_{pmk}-\frac{\hat{\sigma}'_{pmk}}{\sqrt{n}}\ \hat{x}'_{\beta_a L},\ \hat{C}_{pmk}-\frac{\hat{\sigma}'_{pmk}}{\sqrt{n}}\ \hat{x}'_{\beta_a U}\right)$$

ⓑ $\mu=M,$

$$\left(\hat{C}_{pmk}-\frac{\hat{\sigma}_{pmk}}{\sqrt{n}}\ \hat{x}_{\beta_a L},\ \hat{C}_{pmk}-\frac{\hat{\sigma}_{pmk}'}{\sqrt{n}}\ \hat{x}_{\beta_a U}\right)$$

ⓒ $M<\mu<USL,$

$$\left(\hat{C}_{pmk}-\frac{\hat{\sigma}''_{pmk}}{\sqrt{n}}\ \hat{x}''_{\beta_a L},\ \hat{C}_{pmk}-\frac{\hat{\sigma}''_{pmk}}{\sqrt{n}}\ \hat{x}''_{\beta_a U}\right)$$

(1) Acceleration Constant for C_{pmk}

We take the function

$$g(\mu)=g(\mu_{1,}\mu_2)=\frac{(d-|\mu_1-M|)}{3\sqrt{(\mu_2-\mu_1^2)+(\mu_1-T)^2}}.$$

First, consider the case of $\mu=T$. The $a_3=a_4=0$ since $g(\cdot)$ does not depend on μ_3 and μ_4.

ⓐ $LSL<\mu<M,$

$$\hat{a}_1=(d+\overline{X}^*-M)\ \overline{X}^*\ 3S^{*3}+1\ 3S^*,$$
$$\hat{a}_2=-(d+\overline{X}^*-M)\ 6S^{*3}$$

ⓑ $\mu=M,$

$$\hat{a}_1=d\ \overline{X}^*\ 3S^{*3},$$
$$\hat{a}_2=-d\ 6S^{*3}$$

ⓒ $M<\mu<USL,$

$$\hat{a}_1=(d-\overline{X}^*+M)\ \overline{X}^*\ 3S^{*3}-1\ 3S^*,$$
$$\hat{a}_2=-(d-\overline{X}^*+M)\ 6S^{*3}$$

Second, consider the case of $\mu\neq T$. The $a_3=a_4=0$ since $g(\cdot)$does not depend on μ_3 and μ_4.

ⓐ $LSL<\mu<M,$

$$\hat{a}_1=\frac{1}{3\sqrt{S^{*2}+(\overline{X}^*-T)^2}}\left[1+\frac{(d+\overline{X}^*-M)T}{3S^{*2}+(\overline{X}^*-T)^2}\right],$$
$$\hat{a}_2=-\frac{d+\overline{X}^*-M}{6\ S^{*2}+(\overline{X}^*-T)^{2\ \frac{3}{2}}}$$

ⓑ $\mu=M,$

$$\hat{a}_1=\frac{dT}{3\ S^{*2}+(\overline{X}^*-T)^{2\ \frac{3}{2}}},$$
$$\hat{a}_2=-\frac{d}{6\ S^{*2}+(\overline{X}^*-T)^{2\ \frac{3}{2}}}$$

ⓒ $M<\mu<USL,$

$$\hat{a}_1=\frac{1}{3\sqrt{S^{*2}+(\overline{X}^*-T)^2}}\times[(d-\overline{X}^*+M)T3S^{*2}+(\overline{X}^*-T)^2-1],$$
$$\hat{a}_2=-d-\overline{X}^*+M6\ S^{*2}+(\overline{X}^*-T)^{2\ \frac{3}{2}}$$

In two cases, $\hat{\mu}_{ijk}$ are as belows.

<Table 1> Coverage of 95% Lower CL and 90% CI for $C_{pmk}(=1.491)$: $N(50,2^2)$

Sample Size	Bootstrap Method	Coverage of 95% Lower CL	Coverage of 90% CI	Average Length of 90% CI	Standard Deviation of 90% CI
n=10	AN	0.907	0.726	0.8516	0.3378
	SB	0.950*	0.847	1.2421	0.5899
	PB	0.878	0.826	1.1538	0.5091
	BCPB	0.885	0.834	1.1323	0.4390
	STUD	0.943*	0.847	1.1889	0.6660
	HYB	0.974	0.781	1.1538	0.5091
	ABC	0.903	0.829	1.0658	0.3819
n=30	AN	0.952*	0.830	0.5791	0.1322
	SB	0.959*	0.873	0.6315	0.1320
	PB	0.921	0.865	0.6240	0.1287
	BCPB	0.928	0.870	0.6242	0.1210
	STUD	0.952*	0.876*	0.6511	0.1658
	HYB	0.974	0.853	0.6240	0.1287
	ABC	0.942*	0.870	0.6143	0.1187
n=50	AN	0.972	0.836	0.4535	0.0799
	SB	0.975	0.876*	0.4878	0.0799
	PB	0.955*	0.867	0.4848	0.0795
	BCPB	0.952*	0.870	0.4853	0.0769
	STUD	0.966*	0.875	0.5005	0.0928
	HYB	0.982	0.869	0.4848	0.0795
	ABC	0.969	0.862	0.4813	0.0772

$$\hat{\mu}_{111} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^3,$$
$$\hat{\mu}_{112} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2(Y_i-\overline{Y}),$$
$$\hat{\mu}_{122} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})^2,$$
$$\hat{\mu}_{222} = \frac{1}{n}\sum_{i=1}^{n}(Y_i-\overline{Y})^3,\ Y_i=X_i^2,\ \overline{Y}=\frac{1}{n}\sum_{i=1}^{n}X_i^2,$$
$$\hat{\mu}_{121} = \hat{\mu}_{211} = \hat{\mu}_{112},\ \hat{\mu}_{212} = \hat{\mu}_{221} = \hat{\mu}_{122}$$

3. Simulation result for our process capability index C_{pmk}

Comparing some confidence limits, we consider the case when the underlying distribution is normal. And we restrict our discussion that sample size n coincides with bootstrap sample size m(n=m=10, 30, 50). Also, we choose two control limits USL=60 and LSL=40. Considering Franklin and Wasseman(1992)'s design of simulation experiment, we choose two process means 50, 52 and two process variances 2^2, 3^2.

For the various methods discussed in section III, simulation results are tabulated in Tables 1 through 4. First of all, we note the results of the 95% bootstrap lower confidence limits. The limits of SB and STUD are well achieved for the nominal coverage 0.95 of the index C_{pmk}. In fact, most of the practical coverages of SB and STUD confidence intervals are contained in the interval (0.933, 0.967) for the true value of 0.95. On the other hand, the other limits show significant departures from 0.95, especially, lower than the nominal value 0.95. However, as we expect, all of these bootstrap limits tend to increase towards to 0.95 as the sample size n increases but the rate is slow particularly for the PB limits.

Also for the 90% bootstrap confidence intervals, we obtained the result that the limits of SB and STUD are well achieved in proportions consistently near the nominal coverage 0.90 for the index C_{pmk}. Similarly a 99% confidence interval for the coverage proportion of a true 90% confidence interval would be (0.876, 0.924). None of the SB and STUD confidence intervals were outside this range too. Such results seem to validate the simulation (since we may expect the limits as a normal result) and they also validate the performance of the SB and STUD method as being equivalent in coverage performance for the index C_{pmk} under the assumption of an underlying normal process.

<Table 2> Coverage of 95% Lower CL and 90% CI for $C_{pmk}(=1.054)$: $N(50,3^2)$

Sample Size	Bootstrap Method	Coverage of 95% Lower CL	Coverage of 90% CI	Average Length of 90% CI	Standard Deviation of 90% CI
n=10	AN	0.930*	0.674	0.5839	0.2278
	SB	0.973	0.839	0.9204	0.4232
	PB	0.923	0.811	0.8542	0.3662
	BCPB	0.928	0.830	0.8697	0.3447
	STUD	0.959*	0.807	0.8940	0.4680
	HYB	0.982	0.751	0.8542	0.3662
	ABC	0.939*	0.809	0.8085	0.2785
n=30	AN	0.957*	0.812	0.4097	0.0789
	SB	0.967*	0.876*	0.4766	0.1100
	PB	0.939*	0.874	0.4707	0.1080
	BCPB	0.948*	0.883*	0.4756	0.1020
	STUD	0.956*	0.881*	0.4998	0.1532
	HYB	0.951*	0.851	0.4708	0.1080
	ABC	0.952*	0.881*	0.4671	0.1011
n=50	AN	0.968	0.836	0.3233	0.0611
	SB	0.966*	0.881*	0.3617	0.0628
	PB	0.957*	0.873	0.3593	0.0629
	BCPB	0.950*	0.882*	0.3635	0.0607
	STUD	0.956*	0.893*	0.3756	0.0792
	HYB	0.970	0.861	0.3593	0.0629
	ABC	0.959*	0.878*	0.3586	0.0604

<Table 3> Coverage of 95% Lower CL and 90% CI for $C_{pmk}(=1.193)$: $N(52, 2^2)$

Sample Size	Bootstrap Method	Coverage of 95% Lower CL	Coverage of 90% CI	Average Length of 90% CI	Standard Deviation of 90% CI
n=10	AN	0.903	0.797	0.8918	0.2746
	SB	0.960*	0.890*	1.2485	0.5787
	PB	0.868	0.811	1.1509	0.4940
	BCPB	0.879	0.816	1.0715	0.4224
	STUD	0.945*	0.889*	1.1876	0.6527
	HYB	0.979	0.809	1.1509	0.4940
	ABC	0.900	0.825	1.0093	0.3637
n=30	AN	0.935*	0.858	0.5807	0.1209
	SB	0.953*	0.887*	0.6319	0.1457
	PB	0.895	0.849	0.6209	0.1420
	BCPB	0.910	0.863	0.6040	0.1320
	STUD	0.953*	0.982*	0.6528	0.1594
	HYB	0.964*	0.859	0.6209	0.1420
	ABC	0.930	0.864	0.5951	0.1286
n=50	AN	0.946*	0.872	0.4578	0.0717
	SB	0.959*	0.890*	0.4806	0.0804
	PB	0.917	0.883*	0.4759	0.0787
	BCPB	0.932	0.883*	0.4688	0.0758
	STUD	0.963*	0.889*	0.4837	0.0916
	HYB	0.978	0.866	0.4759	0.0787
	ABC	0.947*	0.881*	0.4653	0.0749

<Table 4> Coverage of 95% Lower CL and 90% CI for $C_{pmk}(=0.843)$: $N(52, 3^2)$

Sample Size	Bootstrap Method	Coverage of 95% Lower CL	Coverage of 90% CI	Average Length of 90% CI	Standard Deviation of 90% CI
n=10	AN	0.862	0.785	0.6584	0.2020
	SB	0.928	0.867	0.9131	0.4211
	PB	0.852	0.825	0.8461	0.3602
	BCPB	0.859	0.828	0.8000	0.3091
	STUD	0.951*	0.836	0.8915	0.4627
	HYB	0.958*	0.809	0.8461	0.3602
	ABC	0.880	0.827	0.7581	0.2735
n=30	AN	0.927	0.866	0.4406	0.0884
	SB	0.941*	0.887*	0.4800	0.1069
	PB	0.900	0.870	0.4716	0.1044
	BCPB	0.910	0.870	0.4602	0.0976
	STUD	0.958*	0.893*	0.4993	0.1438
	HYB	0.961*	0.871	0.4716	0.1044
	ABC	0.929	0.877*	0.4552	0.0951
n=50	AN	0.927	0.868	0.3471	0.0523
	SB	0.936*	0.883*	0.3636	0.0583
	PB	0.901	0.866	0.3605	0.0576
	BCPB	0.911	0.871	0.3554	0.0554
	STUD	0.952*	0.894*	0.3728	0.0752
	HYB	0.959*	0.878*	0.3605	0.0576
	ABC	0.927	0.875	0.3533	0.0551

Reference

[1] Bickel,P.J. and Freedman,D.A..(1981). Some Asymptotic Theory for the Bootstrap, Annals of Statistics, 9, 1196-1217.

[2] Chen,S. and Hsu,N.(1995). The Asymptotic Distribution of the Process Capability Index C_{pmk}, Communications in Statistics, Theory and Methods, 24(5), 1279-1291.

[3] Cho,J.J., Kim,J.S. and Park,B.S.(1999). Better Nonparametric Bootstrap Confidence Intervals for Capability Index C_{pk}. The Korean Journal of Applied Statistics, 12(1), 45-65.

[4] Efron,B.(1979). Bootstrap methods: Another look at the jackknife. Annals of Statistics, 9, 139-172.

[5] Efron,B. and Tibshirani,R.(1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. Statistical Science, 1, 54-77.

[6] Franklin,L.A. and Wasserman,G.(1991). Bootstrap Confidence Interval Estimate of C_{pk}: An Introduction. Communications in Statistics-Simulation and Computation, 20(1), 231-242.

[7] Franklin,L.A. and Wasserman,G.(1992). Bootstrap Lower Confidence Interval Limits for Capability Indices. Journal of Quality Thechnology, 24, 196-210.

[8] Gunter,B.H.(1989). The Use and Abuse of C_{pk}, Part 2, Quality Progress, 22(3), 108-109.

[9] Hall,P.(1988). Theoretical Comparison of Bootstrap Confidence Intervals. Annals of Statistics, 16, 927-953.

[10] Mallows,C.L.(1972). A Note on Asymptotic Joint Normality, Annals of Mathematical Statistics, 43, 508-515.

[11] Pearn,W.L., Kotz,S. and Johnson,N.L. (1992). Distributional and inferential properties of process capability indices. Journal of Quality Technology, 24, 216-231.

[12] Wright P.A.(1998). The Probability Density Function of Process Capability Index C_{pmk}, Communications in Statistics, Theory and Methods, 27(7), 1781-1789.