Bootstrap Estimation for the Process Incapability Index C_{pp}

JeongHye Han
Department of Industrial Systems Engineering, Yonsei University,
Seoul, Korea

JoongJae Cho
Department of Statistics, Chungbuk National University,
Chungbuk, Korea

Chun Sung Lim
Department of Industrial Systems Engineering, Yonsei University,
Seoul, Korea

Abstract

Process Capability can be expressed with a process index which indicates the incapability of a process to meet its specifications. This index is regarded as a process capability index(PCI) or more precisely as a process incapability index(PII). It is obtained from a simple transformation of a PCI.

Greenwich and Jahr-Schaffrath(1995) considered the PII C_{pp} which could be obtained from the transformation to the PCI, C'_{pm}, and they provided the asymptotic distribution for C_{pp} which was useful unless the process characteristic was normally distributed.

However, some statistical inferences based on the asymptotic distribution need a large sample size. There are some processes which process engineers could not help obtaining sufficiently a large sample size. Thus, we have derived its corresponding bootstrap asymptotic distribution since bootstrapping would be a helpful technique for the PII, C_{pp} which was nonparametric or free from assumptions of the distribution of the characteristic X.

Moreover, we have constructed six bootstrap confidence intervals used in reducing bias of estimations based on the bootstrap asymptotic distribution and simulated their performances for C_{pp}.

1. Introduction

Process capability can be expressed with a process index which indicates the incapability of a process to meet its specifications. This index is regarded as a PCI or more precisely as a PII which is obtained from a simple transformation of a PCI. The PII contains the same information as PCI. Moreover, the PII provides an uncontaminated separation between information concerning the process accuracy and precision. This kind of information separation is not available with PCI.

Even when the underlying process is assumed to be normal, the computation of a PII results in the estimator being a random variable possessing a very complicated distribution. Thus, there exist many studies based on large sample properties for PCIs such as Chan *et al.*(1990), Chen and Hsu(1995), and Chen and Kotz(1996). But there exist few studies on the asymptotic distributions for PIIs including Greenwich and Jahr-Schaffrath(1995).

Generally the central limit theorem is not of use since it is demanded at least 100 observations. Moreover, since many processes can be frequently skewed or heavy tailed in practice, the interval estimation technique that is free from the assumption of distribution is desirable.

Bootstrap sampling techniques can be successfully used to construct reasonably accurate estimates for the PCIs or PIIs when process information is limited. Bootstrap method is precisely such a technique applied for hundreds of papers for that reason. However, there are no studies on bootstrap asymptotic distributions as well as confidence limits for the PIIs yet.

In this paper, we will derive the bootstrap asymptotic distributions for the PII, C_{pp}, and construct six bootstrap confidence intervals for C_{pp} based on Hall(1988). Also, we will present the simulated results to compare their performances based on the consistency of bootstrap.

2. Process Incapability Index

2.1 Process Incapability Index

Assume that the characteristic variable X is distributed with mean, μ, and standard deviation, σ, where they are not known in practice. An allowable process spread called process specification generally consists of lower and upper specification limits (LSL, USL). A target value T somewhere between these limits, and d=(USL-LSL)/2 that is half the length of the specification interval on the characteristic X of each item.

Now we consider the definition of the PII, C_{pp}. Greenwich and Jahr-Schaffrath (1995) applied of the transformation to C'_{pm} for $T \neq M$.

$$
\begin{aligned}
C_{pp} &= C'^{-2}_{pm} \\
&= \left\{ \frac{3\tau}{\min(USL-T, T-LSL)} \right\}^2 \\
&= \left(\frac{\mu-T}{D} \right)^2 + \left(\frac{\sigma}{D} \right)^2 \\
&= C'_{ia} + C'_{ip},
\end{aligned}
$$

where $D = \min\left(\frac{USL-T}{3}, \frac{T-LSL}{3}\right), C'_{ia} = \left(\frac{\mu-T}{D}\right)^2, C'_{ip} = \left(\frac{\sigma}{D}\right)^2$. Since the denominators of subindices are identical, they provide the relative magnitudes of the contributions to the process in capability indicated by C_{pp}.

2.2 Asymptotic Distribution Result

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 standard deviation σ. For asymptotic properties, consider the natural plug-in estimator of C_{pp} as belows.

$$
\hat{C}_{pp} = \left\{ \frac{3\sqrt{S^2 + (\overline{X}-T)^2}}{\min(USL-T, T-LSL)} \right\}^2
$$

The asymptotic distributions of the estimator for C_{pp} have been appeared in Greenwich and Jahr-Schaffrath(1995) as belows.

Theorem If $\mu_4 = E(X-\mu)^4$ exists , then

(a) $\sqrt{n}(\hat{C}_{pp} - C_{pp}) \xrightarrow{d} \frac{1}{D^2}\{2(\mu-T)Z_1 + Z_2\} \equiv N(0, \sigma^2_{pp})$,

where $\sigma^2_{pp} = \frac{1}{D^4}\{(\mu_4 - \sigma^4) + 4(\mu-T)[\sigma^{2(\mu}-T) + \mu_3]\}$.

(b) $\sqrt{n}(\widehat{C_{ia}} - C_{ia}) \xrightarrow{d} N(0, \sigma^2_{ia})$, *where* $\sigma^2_{ia} = \frac{4(\mu-T)^2\sigma^2}{D^4}$.

(c) $\sqrt{n}(\widehat{C}_{ip}-C_{ip}) \xrightarrow{d} N(0,\sigma^2_{ip})$, where $\sigma^2_{ip}=\dfrac{(\mu_4-\sigma^4)}{D^4}$.

(d) $\sqrt{n}(\widehat{C}_{cop}-C_{cop}) \xrightarrow{d} N(0,\sigma^2_{cop})$,

$$\text{where } \sigma^2_{cop}=\frac{9}{(3D-\mu+T)^4}\left[\frac{\sigma^4}{(3D-\mu+T)^2}+\frac{\mu_3}{(3D-\mu+T)}+\frac{(\mu_4-\sigma^4)}{4\sigma^2}\right], D=\min\left(\frac{USL-T}{3},\frac{T-LSL}{3}\right).$$

Proof. See pp.65~68 of Greenwich and Jahr-Schaffrath(1995). □

3. Bootstrapping PIIs

In this section, we introduce the bootstrap algorithm for deriving asymptotic distributions and confidence limits based on bootstrap.

Let $X_1,X_2,\cdots,X_n$ be a random sample of size n from a possible population with distribution F, $t(X_1,X_2,\cdots,X_n;F)$ be the specified random variable of interest, and F_n be the empirical distribution function of $X_1,X_2,\cdots,X_n$. Putting mass $1/n$ at each of the points $X_1,X_2,\cdots,X_n$, n times, we get the bootstrap sample of size n, $X_1^*,X_2^*,\cdots,X_n^*$. The bootstrap method is to approximate the distribution of $t(X_1,X_2,\cdots,X_n;F)$ under F by that of $t(X_1^*,X_2^*,\cdots,X_n^*;F_n)$. A formal description of the bootstrap algorithm goes as follows.

- Step 1 : Given $\mathrm{x}_n=(X_1,X_2,\cdots,X_n)$, the bootstrap sample of size n, $X_1^*,X_2^*,\cdots,X_n^*$ can be obtained with replacement, which is conditionally independent with common distribution F_n.
- Step 2 : From the bootstrap sample $X_1^*,X_2^*,\cdots,X_n^*$, compute the sample mean $\overline{X}^*$ and sample variance S^{*2}

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

- Step 3 : Compute the bootstrap plug-in estimator of C_{pp},

$$\widehat{C}^*_{pp}=\left\{\frac{3\sqrt{S^{*2}+(\overline{X}^*-T)^2}}{\min(USL-T,T-LSL)}\right\}^2$$

Here, we will derive their corresponding bootstrap asymptotic distributions, which may be useful unless the process characteristic is normally distributed.

Theorem If $\mu_4 = E(X-\mu)^4$ exists , then

(a) $\sqrt{n}(\widehat{C^*_{pp}} - \hat{C}_{pp}) | \mathrm{x}_n \xrightarrow{d} \frac{1}{D^2}\{2(\mu-T)Z_1+Z_2\} \equiv N(0,\sigma^2_{pp}),$

$$where\ \sigma^2_{pp} = \frac{1}{D^4}(\mu_4-\sigma^4)+4(\mu-T)[\sigma^2(\mu-T)+\mu_3].$$

(b) $\sqrt{n}(\widehat{C^*_{ia}} - \hat{C}_{ia}) | \mathrm{x}_n \xrightarrow{d} N(0,\sigma^2_{ia}),\ where\ \sigma^2_{ia} = \frac{4(\mu-T)^2\sigma^2}{D^4}.$

(c) $\sqrt{n}(\widehat{C^*_{ip}} - \hat{C}_{ip}) | \mathrm{x}_n \xrightarrow{d} N(0,\sigma^2_{ip}),\ where\ \sigma^2_{ip} = \frac{(\mu_4-\sigma^4)}{D^4}.$

(d) $\sqrt{n}(\widehat{C^*}_{cop} - \hat{C}_{cop}) | \mathrm{x}_n \xrightarrow{d} N(0,\sigma^2_{cop}),$

$$where\ \sigma^2_{cop} = \frac{9}{(3D-\mu+T)^4}\left[\frac{\sigma^4}{(3D-\mu+T)^2}+\frac{\mu_3}{(3D-\mu+T)}+\frac{(\mu_4-\sigma^4)}{4\sigma^2}\right],\quad D=\min\left(\frac{USL-T}{3},\frac{T-LSL}{3}\right).$$

4. Bootstrap Confidence Limits

In this section we will construct six bootstrap confidence limits for C_{pp}. For constructing the bootstrap confidence limits, consider the empirical distributions based on bootstrap as below.

$$\hat{H}(x_\alpha) = P\left(\frac{\sqrt{n}(\widehat{C^*}_{pp} - \hat{C}_{pp})}{S_{C_{pp}}} \le \hat{x}_\alpha \,\middle|\, \mathrm{x}_n\right) = \alpha$$

$$\hat{K}(y_\alpha) = P\left(\frac{\sqrt{n}(\widehat{C^*}_{pp} - \hat{C}_{pp})}{S^*_{C_{pp}}} \le \hat{y}_\alpha \,\middle|\, \mathrm{x}_n\right) = \alpha,$$

where $\hat{x}_\alpha$ and $\hat{y}_\alpha$ denote α-level quantiles, and $S^{*2}_{C_{pp}}$ implies the bootstrap estimator of the variance $\sigma^2_{C_{pp}}$.

Then, a upper $(1-\alpha)100$ % confidence limit can be constructed by using only the upper limit. That is, we consider the following six bootstrap confidence intervals.

SB method: $\left(0,\quad \hat{C}_{pp} + z_{\alpha} S^{*}_{C_{pp}}\right),$

PB method: $\left(0,\quad \widehat{C^{*}}_{pp}((1-\alpha)B)\right)$

BCPB method: $\left(0,\quad \widehat{C^{*}}_{pp}(P_U B)\right).$

STUD method: $\left(0,\quad \hat{C}_{pp} + \frac{S_{C_{pp}}}{\sqrt{n}}\, \hat{y}_{(1-\alpha)}\right).$

HYB method: $\left(0,\quad \hat{C}_{pp} + \frac{S_{C_{pp}}}{\sqrt{n}}\, \hat{x}_{(1-\alpha)}\right).$

ABC method: $\left(0,\quad \hat{C}_{pp} + \frac{S_{C_{pp}}}{\sqrt{n}}\, \hat{x}_{\widehat{\beta_{\alpha U}}}\right).$

We have given simulated results for the PII, C_{pp} to evaluate which upper confidence intervals are better than the others with small sample. For each simulation study, a sample of size n was drawn and for each of size n, B=1000 bootstrap resamples were drawn from that single sample. This single simulation was then replicated N=1000 times. Thus, we are able to calculate the proportion of covering with various bootstrap upper limits. We also have provided the simulation results under normality, lognormal, and chi-square to consider skewness as well as t-distribution, heavy tail.

5. Conclusion

In this paper we have derived bootstrap asymptotic distributions for the PII, C_{pp}, and studied six bootstrap confidence limits for C_{pp}.

Our experiment results are partly identical to those presented in Franklin and Wasserman(1991, 1992), Kim and Cho(1995), Cho and Park(1998) and Han et al.(1998). In Han et al.(1998) the coverage proportions for the STUD, ABC, BCPB limits based on normality were generally stable and worked well for C_s, but in Cho and Park(1998) SB, STUD, and ABC limits did for C_{pk}. Also, SB and PB method were worst. Particularly, ABC method needed much computing time and memory than others. For non-normal process distribution, STUD, HYB, and ABC method tended to approach the true value.

References

1. Chan,L.K., Xiong,Z. and Zhang,D. "On the Asymptotic Distributions of Some Process Capability Indices", *Communications in Statistics: Theory and Methods*, 19(1), 11-18, 1990.
2. Chen,H. and Kotz,S. "Asymptotic Distribution of Wright's Process Capability Index sensitive to Skewness", *Journal of Statistical Computation and Simulation*, 55, 147-158, 1996.
3. Chen,S. and Hsu,N. "The Asymptotic Distribution of the Process Capability Index C_{pmk}", *Communications in Statistics: Theory and Methods*, 24(5), 1279-1291, 1995.
4. Cho,J.J. and Park,B.S. "Better Nonparametric Bootstrap Confidence Interval for Process Capability Index C_{pk}", A manuscript submitted for publication, 1998.
5. Franklin, L.A. and Wasserman, G.S. "Bootstrap Confidence Interval Estimates of C_{pk}", *Communications in Statistics:Simulation and Computation*, 20(1), 231-242, 1991.
6. Franklin, L.A. and Wasserman, G.S. "Bootstrap Lower Confidence Interval Limits for Capability Indices", *Journal of Quality technology*, 24, 196-210, 1992.
7. Greenwich, M. and Jahr-Schaffrath, B.L. "A Process Incapability Index", *International Journal of Quality & Reliability Management*, 12(4), 58-71, 1995.
8. Hall,P. "Theoretical Comparison of Bootstrap Confidence Intervals", *Annals of Statistics*, 16, 927-953, 1988.
9. Han,J.H., Cho,J.J, and Leem,C.S. "Bootstrap Confidence Limits for Wright's $C_s(\gamma)$", A manuscript submitted for publication, 1998.
10. Kim,P.K. and Cho,J.J. "Bootstrapping Some Process Capability Indices", *Journal of the Korean Society for Quality Management*, 23(4), 157-166, 1995.