대한산업공학회/한국경영과학회 2000 춘계공동학술대회 논문집. 2000년4월21-22일. 경남대학교
Session OC5.6

비대칭 모집단에 대한 공정능력지수의 개발
Process Capability Indices for Skewed Populations

장영순, 배도선
한국과학기술원 산업공학과

Abstract

This paper proposes a new heuristic method of constructing process capability indices (PCIs) for skewed populations. It is based on weighted standard deviation (WSD) method which decomposes the standard deviation of a quality characteristic into upper and lower deviations and adjusts the value of PCI using decomposed deviations in accordance with the skewness estimated from sample data. For symmetric populations, the proposed PCIs reduce to standard PCIs. Asymptotic distributions of the estimators of the PCIs are obtained. The performances of the proposed methods are compared with those of the standard and other methods. Numerical comparisons indicate that considerable improvements over existing methods can be achieved by the use of WSD method when the underlying distribution is skewed.

1. Introduction

Process capability analysis is an important and integrated part of the statistical process control activities for the continuous improvement of quality and productivity. The capability of a process is frequently measured by a process capability index (PCI) which is a dimensionless function of process parameters and specifications. PCI is designed to provide a common and easily understood language for quantifying the performance of a process. The standard PCIs such as C_p and C_{pk} are usually determined under the assumption that the quality characteristic follows a normal distribution. In many situations, however, we may have reason to doubt the validity of the normality assumption. For example, the measurements from drilling, coating, chemical, and semiconductor processes often follow skewed distribution.

For skewed populations, the proportion of nonconforming items for the fixed value of standard PCIs tend to increase as skewness increases. The trouble arises from the fact that the standard PCIs ignore the skewness of the underlying population. Therefore, a method of adjusting the values of PCI in accordance with the expected proportion of NC items by considering the skewness of the underlying population is desirable. Some authors have addressed this problem. See Rodriguez (1992), English and Taylor (1993), Pyzdek (1995), Somerville and Montgomery (1996-97), Kotz and Lovelace (1998), and Shore (1998).

This paper proposes a new heuristic method of constructing simple PCIs for skewed populations based on the weighted standard deviation (WSD) method of Chang and Bai (2000). This method adjusts the values of PCIs in accordance with the degree of skewness of the underlying population by using different factors in computing the deviations above and below the process mean. When the underlying population is symmetric, however, these indices reduce to standard PCIs. The proposed PCIs are easily calculated with hand-held calculator from small samples and perform better than PCIs based on WV method of Choi and Bai (Kotz and Lovelace, 1998, Ch.4.6) and Clements' method (Clements, 1989). We also derive the asymptotic distributions of estimators of proposed PCIs.

2. PCIs Based on the WSD Method

2.1 WSD Method

Chang and Bai (2000) proposed the WSD

method of constructing $\overline{X}$ and CUSUM control charts for skewed populations. The WSD method is based on the idea that an asymmetric distribution can be divided into two parts at its mean and each part can be treated as a half of a symmetric distribution, each having the same means but different standard deviations. Based on the ratios of standard deviations of the two symmetric distributions, standard deviation σ can be decomposed into upper and lower deviations $WSD_U = P\sigma$ and $WSD_L = (1-P)\sigma$, where $P = \Pr\{X \le \mu\}$. For $\overline{X}$ charts, $2WSD_U$ and $2WSD_L$ are used for upper and lower control limits, respectively. See Chang and Bai (2000) for detailed discussions on the basic idea. The WSD method can be used to adjust the values of PCIs according to the degree of skewness of the underlying distribution by using different deviations in computing the upper and lower PCIs. That is, $2WSD_U$ and $2WSD_L$ can be used to construct PCIs for skewed populations.

2.2 C_p based on the WSD method

The C_p based on the WSD method, C_p^{WSD}, is defined as

$$\begin{aligned} C_p^{WSD} &= \min\left\{\frac{U-L}{6\cdot 2WSD_U}, \frac{USL-LSL}{6\cdot 2WSD_L}\right\} \\ &= \frac{U-L}{6\sigma}\min\left\{\frac{1}{2P}, \frac{1}{2(1-P)}\right\} \\ &= C_p/D \ , \end{aligned}$$

where U and L are upper and lower specification limits, respectively, and $D = 1 + |1-2P|$. In formula (1), $2WSD_U$ and $2WSD_L$ are used in place of σ to reflect the degree of skewness. If the underlying distribution is symmetric, $P=0.5$ and $D=1.0$, i.e., $C_p^{WSD} = C_p$. However, if it is skewed, $D > 1.0$ and $C_p^{WSD} < C_p$.

Table 1. C_P^{WSD} for normal and lognormal distributions with $\mu=40$ and $\sigma=10$

Process	α_3	P	ENC[1]	C_p^{WSD}	C_p
A	0	0.50	2,700	**1.00**	1.00
B	1	0.56	10,461	**0.89**	1.00
C	2	0.61	16,358	**0.82**	1.00
D	3	0.64	18,325	**0.78**	1.00

[1] ENC: expected number of nonconforming items per million

Table 1 represents the effect of C_p^{WSD} for skewed populations. The underlying distribution of process A is normal and the distributions of B, C, and D are three parameter lognormals with skewness $\alpha_3 = 1, 2,$ and 3, respectively, and have common mean $\mu=40$ and standard deviation $\sigma=10$. The specification limits are $U=70$ and $L=10$, and hence $C_p=1.0$ for processes A, B, C, and D. We note that the proportion of nonconforming items increases as skewness increases and C_p^{WSD} represents such a phenomenon, i.e., C_p^{WSD} decreases as skewness increases whereas C_p remains constant.

2.3 C_{pk} based on the WSD method

Consider the situation where only a single specification limit exists. Under the WSD method, upper and lower capability indices are defined as

$$C_{pku}^{WSD} = \frac{U-\mu}{3\cdot 2WSD_U} = \frac{U-\mu}{6P\sigma}\ ,$$
$$C_{pkl}^{WSD} = \frac{\mu-L}{3\cdot 2WSD_L} = \frac{\mu-L}{6(1-P)\sigma}\ .$$

The PCI for the case of two-side specifications is

$$C_{pk}^{WSD} = \min\{C_{pku}^{WSD}, C_{pkl}^{WSD}\}.$$

Note that for symmetric population, C_{pk}^{WSD} reduces to C_{pk}, and if μ is at the center of the two specification limits, C_{pk}^{WSD} reduces to C_p^{WSD}.

Table 2. C_{pk}^{WSD} for lognormal distributions with $\sigma=10$ and $\alpha_3=2$

Process	μ	ENC	C_{pk}^{WSD}	C_{pk}	C_p^{WSD}	C_p
A	40	16,358	**0.82**	1.00	0.82	1.00
B	35	10,400	**0.96**	0.83	0.82	1.00
C	45	26,809	**0.68**	0.83	0.82	1.00

Table 2 shows C_{pk}^{WSD} and other PCIs for three lognormal distributions having the same shape and scale, but different locations. The means of process A, B, and C are $\mu=40$, 35, and 45, respectively, and have common $\sigma=10$ and $\alpha_3=2$. The upper and lower specification limits are $U=70$ and $L=10$. Note that, in Table 2, C_{pk}^{WSD} for process B, which produces the least nonconforming items, is larger than that for process A or C, while C_{pk}

indicates that process A is most capable, and C_p and C_p^{WSD} remain constant. That is, C_{pk}^{WSD} can describe the capability of the process more accurately when the underlying distribution is skewed and the process is not centered.

2.4 Performance of the proposed PCIs

In this section, C_p^{WSD} is numerically compared with C_p' using Clements' method and C_p^{WV} based on WV method when the underlying distributions are Weibull, lognormal, and gamma.

Table 3. ENC, C_p^{WSD}, C_p^{WV}, C_p', and matched C_p for Weibull, lognormal, and gamma distribution

Distri-bution	α_3	ENC	C_p^{WSD}	C_p^{WV}	C_p'	matched C_p
Weibull	0.5	4,227	**0.94**	0.97	1.10	0.95
	1.0	9,870	**0.88**	0.94	1.04	0.86
	1.5	14,915	**0.83**	0.91	0.98	0.81
	2.0	18,316	**0.79**	0.89	0.91	0.79
	2.5	20,28,	**0.76**	0.87	.	0.77
	3.0	21,256	**0.74**	0.86	.	0.76
lognormal	0.5	5,639	**0.94**	0.97	0.98	0.92
	1.0	10,461	**0.89**	0.94	0.94	0.85
	1.5	14,087	**0.85**	0.92	0.89	0.82
	2.0	16,358	**0.82**	0.91	0.84	0.80
	2.5	17,653	**0.80**	0.89	.	0.79
	3.0	18,325	**0.78**	0.88	.	0.79
gamma	0.5	5,431	**0.94**	0.97	0.99	0.93
	1.0	10,336	**0.88**	0.94	0.98	0.86
	1.5	14,782	**0.83**	0.91	0.96	0.81
	2.0	18,316	**0.79**	0.89	0.91	0.79
	2.5	20,856	**0.75**	0.87	.	0.77
	3.0	22,528	**0.72**	0.85	.	0.76

· : not applicable

Table 3 gives ENC, C_p^{WSD}, C_p^{WV}, and C_p' as skewness increases. In each case, it is assumed that $U=3$ and $L=-3$, and the distribution is standardized, i.e., shifted and scaled to produce the same value of $\mu=0$ and $\sigma=1$ and hence $C_p=1.0$ for all cases. It also gives the 'matched C_p' calculated by $-(1/3)\,\Phi^{-1}((1/2)\,\mathrm{ENC}\times 10^{-6})$ since $ENC\times 10^{-6}=2\Phi(-3C_p)$ based on normal distribution theory, where $\Phi(\cdot)$ is the cumulative standard normal distribution function. If the value of the PCI is close to that of matched C_p, the PCI can be considered to describe the process capability very well in terms of NC proportion.

Table 3 shows that the values of C_p^{WSD}, C_p^{WV}, and C_p' decrease as skewness and ENC increase. However C_p^{WSD} is very close to the matched C_p regardless of skewness and describes the process better than C_p^{WV} and C_p'.

2.5 Estimation of the proposed PCIs

To use PCIs based on the WSD method in practice, μ, σ, and P must be estimated. We suppose that a random sample $X_1,\cdots,X_n$ of size n is available. Then μ and σ can be estimated by grand mean $\overline{X}_n$ and sample standard deviation S_n. Since P is the probability that X will be less than or equal to μ, it can be estimated by using the number of observations less than or equal to $\overline{X}_n$;

$$\hat{P}_n=\frac{1}{n}\sum_{i=1}^{n} I(\overline{X}_n-X_i),$$

where $I(x)=1$ for $x\ge 0$ and $I(x)=0$ for $x<0$.

The proposed PCIs can be estimated as

$$\hat{C}_p^{\,WSD}=\frac{U-L}{6\,\hat{D}_n S_n}$$

and

$$\hat{C}_{pk}^{\,WSD}=\min\left\{\frac{USL-\overline{X}}{6\,\hat{P}_n S_n},\ \frac{\overline{X}-LSL}{6(1-\hat{P}_n)S_n}\right\},$$

where $\hat{D}_n=1+|1-2\hat{P}_n|$.

3. Asymptotic Properties

We assume that random variable X representing process quality characteristic is continuous with pdf $f(x)$. Let '$\xrightarrow{L}$' and '$\xrightarrow{P}$' denote 'convergence in law' and 'convergence in probability', respectively. Let $N(\mu,\sigma^2)$ 'normal distribution with mean μ and variance σ^2' and $\mu_k=E(X-\mu)^k$ be the k^{th} central moment of X.

[Theorem 1]

(a) $\hat{P}_n\xrightarrow{P}P$,

(b) $\sqrt{n}(\hat{P}_n-P)\xrightarrow{L}Z_1$,

where $Z_1\sim N(0,P(1-P))$.

[Lemma 1]

If μ_4 exists, then

$$\sqrt{n}(\hat{P}_n - P, \overline{X}_n - \mu, S_n^2 - \sigma^2) \xrightarrow{L} (Z_1, Z_2, Z_3),$$

where $(Z_1, Z_2, Z_3) \sim N(0, \Sigma_3)$,

$$\Sigma_3 = \begin{bmatrix} P(1-P) & \mu_a & \sigma_a^2 - P\sigma^2 \\ \mu_a & \sigma^2 & \mu_3 \\ \sigma_a^2 - P\sigma^2 & \mu_3 & \mu_4 - \sigma^4 \end{bmatrix},$$

$0 = (0,0,0)^T$, $\mu_a = \int_{-\infty}^{\mu}(x-\mu)f(x)dx$, and $\sigma_a^2 = \int_{-\infty}^{\mu}(x-\mu)^2 f(x)dx$.

[Lemma 2]

$\hat{C}_p^{WSD}$ and $\hat{C}_{pk}^{WSD}$ are consistent estimators of C_p^{WSD} and C_{pk}^{WSD}, respectively.

[Theorem 2]

If μ_4 exists, then

(a) $\sqrt{n}(\hat{C}_p^{WSD} - C_p^{WSD})$

$$\xrightarrow{L} \begin{cases} V_1 & , \text{ if } P<1/2 \\ -\frac{U-L}{3\sigma}|Z_1| - \frac{U-L}{12\sigma^3}Z_3 & , \text{ if } P=1/2 \\ V_2 & , \text{ if } P>1/2 \end{cases}$$

(b)

$\sqrt{n}(\hat{C}_p^{WSD} - C_p^{WSD})$

$$\xrightarrow{L} \begin{cases} V_3 & , \text{ if } \frac{\mu-L}{1-P} < \frac{U-\mu}{P} \\ \frac{1}{12}\Big\{\frac{(2P-1)(U-L)}{P(1-P)\sigma}Z_1 \\ \quad + \frac{2P-1}{P(1-P)\sigma}Z_2 - \frac{U-L}{\sigma^3}Z_3 & , \text{ if } \frac{\mu-L}{1-P} = \frac{U-\mu}{P} \\ \quad - \Big| \frac{U-L}{P(1-P)\sigma}Z_1 - \frac{1}{P(1-P)\sigma}Z_2 \Big| \Big\} \\ V_4 & , \text{ if } \frac{\mu-L}{1-P} > \frac{U-\mu}{P} \end{cases}$$

where $V_i \sim N(0, \sigma_{V_i}^2)$, $i = 1,2,3,4$,

$$\sigma_{V_1}^2 = \frac{(U-L)^2}{144(1-P)^2\sigma^2}\left\{\frac{2P}{1-P} - \frac{\sigma_a^2}{(1-P)\sigma^2} + \frac{\alpha_4}{4} - \frac{1}{4}\right\},$$

$$\sigma_{V_2}^2 = \frac{(U-L)^2}{144P^2\sigma^2}\left\{\frac{1-P}{P} + \frac{\sigma_a^2}{P\sigma^2} + \frac{\alpha_4}{4} - \frac{5}{4}\right\},$$

$$\sigma_{V_3}^2 = \frac{(\mu-L)^2}{36(1-P)^2\sigma^2}\left\{\frac{2P}{1-P} - \frac{\sigma_a^2}{(1-P)\sigma^2} + \frac{\alpha_4}{4} - \frac{1}{4}\right\} + \frac{\mu-L}{36(1-P)^2\sigma}\left\{\frac{2\mu_a}{(1-P)\sigma} - \alpha_3\right\} + \frac{1}{36(1-P)^2},$$

$$\sigma_{V_4}^2 = \frac{(U-\mu)^2}{36P^2\sigma^2}\left\{\frac{1-P}{P} + \frac{\sigma_a^2}{P\sigma^2} + \frac{\alpha_4}{4} - \frac{5}{4}\right\} + \frac{U-\mu}{36P^2\sigma}\left\{\frac{2\mu_a}{P\sigma} + \alpha_3\right\} + \frac{1}{36P^2},$$

$\alpha_3 = \mu_3/\sigma^3$, and $\alpha_4 = \mu_4/\sigma^4$.

4. Conclusions

We have proposed a simple heuristic method of constructing process capability indices for an arbitrary skewed population based on the weighted standard deviation method. It is shown that the performances of proposed PCIs are better than those of Clements' and WV methods as well as the standard method. Methods of estimating the proposed PCIs are provided, and large sample properties of their estimators are established.

References

Chang, Y. S. and Bai, D. S. (2000), $\overline{X}$ and CUSUM control charts for skewed populations with weighted standard deviations, *manuscript.*

Clements, J. A. (1989), Process capability calculations for nonnormal calculations, *Quality Progress*, **22**, 49-55.

English, J. R. and Taylor, G. D. (1993), Process capability analysis - A robustness study, *International Journal of Production Research*, **31**, 1621-1635.

Kotz, S. and Lovelace, C. R. (1998), *Process capability indices in theory and practice*, Arnold, New York.

Pyzdek, T. (1995), Why normal distributions aren't - All that normal, *Quality Engineering*, **7**, 769-777.

Rodriguez, R. N. (1992), Recent developments in process capability analysis, *Journal of Quality Technology*, **24**, 176-187.

Shore, H. (1998), A new approach to analyzing non-normal quality data with application to process capability analysis, *International Journal of Production Research*, **36**, 1917-1933.

Somerville, S. E. and Montgomery, D. C. (1996-97), Process capability indices and non-normal distributions, *Quality Engineering*, **9**, 305-316.