ESTIMATING PROCESS CAPABILITY INDICES FOR NON-NORMAL PEARSONIAN POPULATIONS

W. L. PEARN AND K. S. CHEN

Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu, Taiwan R.O.C.

SUMMARY

Clements[1] introduced a method for calculating the estimators of two classical process capability indices (PCI), C_p and C_{pk}, for non-normal Pearsonian populations. Pearn and Kotz[2] applied Clements' method to obtain first approximation estimators of PCIs for non-normal populations to the two more advanced PCIs, C_{pm} and C_{pmk} developed by Chan *et al.*[3] and Pearn *et al.*[4] In this paper, we consider a different approach for calculating the estimators of the four PCIs. The new approach may be viewed as a modification of Clements' method. Comparisons between Clements' and the proposed new methods are also provided.

KEY WORDS: capability indices;: non-normal Pearsonian populations; uncentred target

1. INTRODUCTION

Process capability indices (PCIs), whose purpose is to determine whether a manufacturing process is capable of producing items within a specified tolerance, have received substantial research attention in recent years. Several capability indices have been proposed to assess process performance. Examples include the two widely used indices in manufacturing industries, C_p and C_{pk}, and the two more advanced indices, C_{pm} and C_{pmk}. Discussions and analysis of these indices on estimation and construction of confidence intervals have been the focus of many statisticians and quality researchers (see References 3–6, and many others). Most of the investigations, however, depend heavily on the assumption of normal variability.

In a pioneering paper, Clements[1] proposed a method for calculating the estimators of two classical process capability indices (PCI), C_p and C_{pk}, for non-normal Pearsonian populations. The method is essentially based on a set of available sample data for a well in-control process using estimates of the mean, standard deviation, skewness and kurtosis. Under the assumption that these four parameters determine the type of the Pearson distribution curve, Clements[1] used the tables provided by Gruska *et al.*[7] for percentages of the family of Pearson curves as a function of skewness and kurtosis. In this paper, we investigate Clements' method, and propose another approach for calculating the estimators of PCIs. The new method may be viewed as a modification of Clements'. Numerical examples are provided to compare Clements' and the proposed new methods.

2. CLEMENTS' METHOD

To estimate the value of the index $C_p = (\mathrm{USL} - \mathrm{LSL})/6\sigma$, where USL and LSL are upper and lower specification limits, and σ is the standard deviation of the process, Clements replaced 6σ by $U_p - L_p$ where U_p is the 99·865 percentile and L_p is the 0·135 percentile determined from Gruska *et al.*'s table[7] for the particular values of skewness and kurtosis which are calculated from the sample data. The idea behind such replacements is to mimic the property of the normal distribution for which the tail probability outside the $\pm 3\sigma$ limits from μ is 0·27 per cent thus ensuring that if the calculated value of $C_p = 1$ (assuming that the process is well-centred) the probability that the process is outside the specification limits (LSL, USL) will be negligibly small. The same approach is used for the index $C_{pk} = \text{minimum } \{(\mathrm{USL} - \mu)/3\sigma, (\mu - \mathrm{LSL})/3\sigma\}$ where μ is the process mean estimated by the median, M, and the two 3σ are estimated by $U_p - M$ and $M - L_p$ for the right-hand and left-hand sides, respectively. Clements' estimators for C_p and C_{pk} are thus defined as

$$\hat{C}_p = \frac{\mathrm{USL} - \mathrm{LSL}}{U_p - L_p}$$

$$\hat{C}_{pk} = \text{minimum}\left\{\frac{\mathrm{USL} - M}{U_p - M}, \frac{M - \mathrm{LSL}}{M - L_p}\right\}$$

Pearn and Kotz[2] applied Clements' method to obtain estimators of process capabilities for non-normal Pearsonian populations to the two more advanced capability indices, C_{pm}[3] and C_{pmk}.[4] Those estimators are

$$\hat{C}_{pm} = \frac{\mathrm{USL} - \mathrm{LSL}}{6\sqrt{\{[(U_p - L_p)/6]^2 + (M - T)^2\}}}$$

$$\hat{C}_{pmk} = \text{minimum}\left\{\frac{\mathrm{USL} - M}{3\sqrt{\{[(U_p - M)/3]^2 + (M - T)^2\}}}, \frac{M - \mathrm{LSL}}{3\sqrt{\{[(M - L_p)/3]^2 + (M - T)^2\}}}\right\}$$

The corresponding Vännman's superstructure[8] for the above four estimators may be written as

$$\hat{C}_p(u, v) = (1 - u)\frac{\mathrm{USL} - \mathrm{LSL}}{6\sqrt{\{[(U_p - L_p)/6]^2 + v(M - T)^2\}}} + u \times \min\left\{\frac{\mathrm{USL} - M}{3\sqrt{\{[(U_p - M)/3]^2 + v(M - T)^2\}}}, \frac{M - \mathrm{LSL}}{3\sqrt{\{[(M - L_p)/3]^2 + v(M - T)^2\}}}\right\}$$

It is easy to verify that

$$\hat{C}_p(0,0) = \hat{C}_p,\ \hat{C}_p(1,0) = \hat{C}_{pk}$$
$$\hat{C}_p(0,1) = \hat{C}_{pm},\ \hat{C}_p(1,1) = \hat{C}_{pmk}$$

Although cases with a centred target ($T = (\mathrm{USL} + \mathrm{LSL})/2$) are quite common in practical situations, there are other situations in which the target does not fall on the midpoint of the specification interval (the target is uncentred). For such cases, Vännman's superstructure[8] can be easily generalized to the following:

$$\hat{C}_p^*(u, v) = (1 - u)\frac{\min\{\mathrm{USL} - T, T - \mathrm{LSL}\}}{3\sqrt{\{[(U_p + L_p)/6]^2 + v(M - T)^2\}}} + u \times \min\left\{\frac{(\mathrm{USL} - T) - |M - T|}{3\sqrt{\{[(U_p - M)/3]^2 + v(M - T)^2\}}}, \frac{(T - \mathrm{LSL}) - |M - T|}{3\sqrt{\{[(M - L_p)/3]^2 + v(M - T)^2\}}}\right\}$$

Consequently, by setting $u = 0$ and 1, $v = 0$ and 1, we obtain the following four estimators for the indices, C_p, C_{pk}, C_{pm} and C_{pmk}, respectively:

$$\hat{C}^*_{\mathrm{p}}(0,0) = \frac{\text{minimum}\,\{\mathrm{USL} - T, T - \mathrm{LSL}\}}{(U_{\mathrm{p}} - L_{\mathrm{p}})/2}$$

$$\hat{C}^*_{\mathrm{p}}(1,0) = \text{minimum}\left\{\frac{(\mathrm{USL} - T) - |M - T|}{U_{\mathrm{p}} - M}, \frac{(T - \mathrm{LSL}) - |M - T|}{M - L_{\mathrm{p}}}\right\}$$

$$\hat{C}^*_{\mathrm{p}}(0,1) = \frac{\text{minimum}\,\{\mathrm{USL} - T, T - \mathrm{LSL}\}}{3\sqrt{\{[(U_{\mathrm{p}} - L_{\mathrm{p}})/6]^2 + (M - T)^2\}}}$$

$$\hat{C}^*_{\mathrm{p}}(1,1) = \text{minimum}\left\{\frac{(\mathrm{USL} - T) - |M - T|}{3\sqrt{\{[(U_{\mathrm{p}} - M)/3]^2 + (M - T)^2\}}}, \frac{(T - \mathrm{LSL}) - |M - T|}{3\sqrt{\{[(M - L_{\mathrm{p}}/3]^2 + (M - T)^2\}}}\right\}$$

3. A NEW ESTIMATING METHOD

In this section, we consider a new estimation method to obtain estimators of C_{p} and C_{pk} for non-normal Pearsonian populations. Instead of estimating the two 3σ by $U_{\mathrm{p}} - M$ and $M - L_{\mathrm{p}}$, respectively, we replace the two 3σ by $(U_{\mathrm{p}} - L_{\mathrm{p}})/2$. The new estimators of C_{p} and C_{pk} can be written as

$$\tilde{C}_{\mathrm{p}} = \frac{\mathrm{USL} - \mathrm{LSL}}{U_{\mathrm{p}} - L_{\mathrm{p}}}$$

$$\tilde{C}_{\mathrm{pk}} = \frac{\min\{\mathrm{USL} - M, M - \mathrm{LSL}\}}{(U_{\mathrm{p}} - L_{\mathrm{p}})/2}$$

Applying the same method to the two more advanced (second and third generations) of PCIs, C_{pm} and C_{pmk}, we obtain

$$\tilde{C}_{\mathrm{pm}} = \frac{\mathrm{USL} - \mathrm{LSL}}{6\sqrt{\{[(U_{\mathrm{p}} - L_{\mathrm{p}})/6]^2 + (M - T)^2\}}}$$

$$\tilde{C}_{\mathrm{pmk}} = \frac{\text{minimum}\{\mathrm{USL} - M, M - \mathrm{LSL}\}}{3\sqrt{\{[(U_{\mathrm{p}} - L_{\mathrm{p}})/6]^2 + (M - T)^2\}}}$$

The corresponding Vännman's superstructure[8] of these new estimators follows immediately:

$$\tilde{C}_{\mathrm{p}}(u, v) = \frac{d - u|M - m|}{6\sqrt{\{[(U_{\mathrm{p}} - L_{\mathrm{p}})/6]^2 + v\,(M - T)^2\}}}$$

It is easy to verify that

$$\tilde{C}_{\mathrm{p}}(0,0) = \tilde{C}_{\mathrm{p}},\ \tilde{C}_{\mathrm{p}}(1,0) = \tilde{C}_{\mathrm{pk}}$$
$$\tilde{C}_{\mathrm{p}}(0,1) = \tilde{C}_{\mathrm{pm}},\ \tilde{C}_{\mathrm{p}}(1,1) = \tilde{C}_{\mathrm{pmk}}$$

In the case where the target is uncentred, Vännman's superstructure can be easily generalized to the following:

$$\tilde{C}^*_{\mathrm{p}}(u, v) = \frac{\min\{\mathrm{USL} - T, T - \mathrm{LSL}\} - u|M - T|}{3\sqrt{\{[(U_{\mathrm{p}} - L_{\mathrm{p}})/6]^2 + v\,(M - T)^2\}}}$$

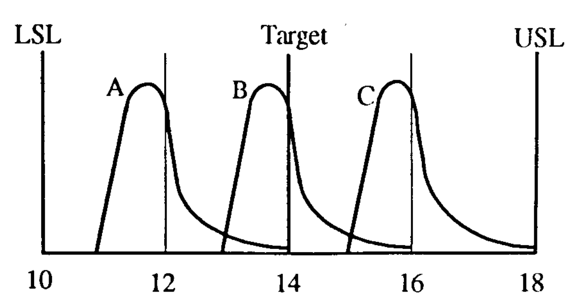
Consequently, by setting $u = 0$ and 1, $v = 0$ and 1, we obtain the following four estimators for the indices, C_{p}, C_{pk}, C_{pm} and C_{pmk}, respectively:

$$\tilde{C}^*_{\mathrm{p}} = \tilde{C}^*_{\mathrm{p}}(0,0) = \frac{\text{minimum}\,\{\mathrm{USL} - T, T - \mathrm{LSL}\}}{(U_{\mathrm{p}} - L_{\mathrm{p}})/2}$$

$$\tilde{C}^*_{\mathrm{pk}} = \tilde{C}^*_{\mathrm{p}}(1,0) = \frac{\text{minimum}\,\{\mathrm{USL} - T, T - \mathrm{LSL}\} - |M - T|}{(U_{\mathrm{p}} - L_{\mathrm{p}})/2}$$

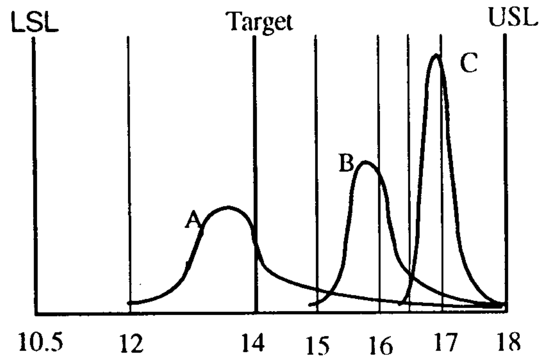
$$\tilde{C}^*_{\mathrm{pm}} = \tilde{C}^*_{\mathrm{p}}(0,1) = \frac{\text{minimum}\,\{\mathrm{USL} - T, T - \mathrm{LSL}\}}{3\sqrt{\{[(U_{\mathrm{p}} - L_{\mathrm{p}})/6]^2 + (M - T)^2\}}}$$

$$\tilde{C}^*_{\mathrm{pmk}} = \tilde{C}^*_{\mathrm{p}}(1,1) = \frac{\text{minimum}\,\{\mathrm{USL} - T, T - \mathrm{LSL}\} - |M - T|}{3\sqrt{\{[(U_{\mathrm{p}} - L_{\mathrm{p}})/6]^2 + (M - T)^2\}}}$$



Process	U_{p}	L_{p}	M	$\hat{C}_{\mathrm{pk}}$	$\tilde{C}_{\mathrm{pk}}$	$\hat{C}_{\mathrm{pmk}}$	$\tilde{C}_{\mathrm{pmk}}$
A	14·00	11·00	12·00	2·00	1·33	0·33	0·32
B	16·00	13·00	14·00	2·00	2·67	2·00	2·67
C	18·00	15·00	16·00	1·00	1·33	0·32	0·32

Figure 1. An example of non-normal populations with a centered target



Process	U_p	L_p	M	$\hat{C}^*_{pk}$	$\tilde{C}^*_{pk}$	$\hat{C}^*_{pmk}$	$\tilde{C}^*_{pmk}$
A	18·00	12·00	14·00	1·00	1·17	1·00	1·17
B	18·00	15·00	16·00	1·00	1·00	0·25	0·24
C	18·00	16·50	17·00	1·00	0·67	0·06	0·06

Figure 2. An example of non-normal populations with an unentered target

4. COMPARISONS

To compare the new estimating method with Clements', we consider the examples depicted in Figure 1 and Figure 2. Figure 1 presents an example of three different non-normal populations with a centred target. The upper specification USL = 18, the lower specification LSL = 10. The target value T for this particular product is preset ot 14. Figure 2 presents an example of three different non-normal populations with an uncentred target. The upper specification USL = 18, the lower specification LSL = 10·5. The target value T for this particular product is also preset to 14. The worksheet provided by Clements (Figure 3 of Reference 1) for calculating the estimators of the capability indices may be used to compute the values of those indices.

In Figure 1 we note that all three processes have same variabilities. Therefore, the quality of process B (which is on-target) is considered to be better than those of processes A and C (which are off-target). Clements' estimator, $\hat{C}_{pk}$, in this case, shows no sensitivity to the target at all ($\hat{C}_{pk} = 2{\cdot}00$ for processes A and B). But, the new estimator, $\tilde{C}_{pk}$, clearly differentiates process B (which is on-target) from processes A and C (which are off-target).

In Figure 2, the quality of process A is considered to be better than that of process B. Similarly, the quality of process B is considered to be better than that of process C. Clements' estimator, $\hat{C}_{pk}$, once again, shows no sensitivity to the target at all in this particular case ($\hat{C}_{pk} = 1{\cdot}00$ for all three processes). But, the new estimator, $\tilde{C}_{pk}$, clearly differentiates process A (which is on-target) from processes B and C (which are off-target).

5. CONCLUSIONS

In this paper, we first investigated Clements' method for calculating the estimators of the four capability indices, C_p, C_{pk}, C_{pm} and C_{pmk} for non-normal populations. Then, we considered a new estimating method to calculate estimators of the four capability indices for non-normal Pearsonian populations. Both cases with centred and uncentred targets are investigated. Superstructures for those estimators were also obtained for centred and uncentred cases. The analysis showed that the estimators calculated from the proposed new method can differentiate on-target processes from off-target processes better than those obtained by applying Clements' method.

REFERENCES

1. J. A. Clements, 'Process capability calculations for non-normal distributions', *Quality Progress*, **22**, 95–100 (1989).
2. W. L. Pearn and S. Kotz, 'Application of Clements' method for calculating second and third generation process capability indices for non-normal Pearsonian populations', to appear in *Quality Engineering*.
3. L. K. Chan, S. W. Cheng and F. A. Spiring, 'A new measure of process capability C_{pm}', *Journal of Quality Technology*, **20**, (3), 162–175 (1988).
4. W. L. Pearn, S. Kotz and N. L. Johnson, 'Distributional and inferential properties of process capability indices', *Journal of Quality Technology*, **24**, (4), 216–231 (1992).
5. V. E. Kane, 'Process capability indices', *Journal of Quality Technology*, **18**, (1), 41–52 (1986).
6. S. Kotz, W. L. Pearn and N. L. Johnson, 'Some process capability indices are more reliable than one might think', *Journal of the Royal Statistical Society, Series C: Applied Statistics*, **42**, (1), 55–62 (1993).
7. G. F. Gruska, K. Mirkhani and L. R. Lamberson, *Non Normal Data Analysis*, Multiface Publishing Co., Michigan, 1989.
8. K. Vännman, 'A unified approach to capability indices', to appear in *Statistica Sinica*.

Authors' biographies:

W. L. Pearn received his Ph.D. degree in Operations Research from the University of Maryland at College Park. He worked for AT&T Bell Laboratories Switch Network Control and Process Quality Centers. Currently, Dr. Pearn is a professor of Operations Research at the Department of Industrial Engineering & Management, National Chiao Tung University, Taiwan, Republic of China.

K. S. Chen received his M.S. degree in Statistics from National Cheng Kung University. Currently, he is a Ph.D. candidate at the Department of Industrial Engineering and Management, National Chiao Tung University.