Int J Adv Manuf Technol (2003) 21:249–256
Ownership and Copyright

The International Journal of Advanced Manufacturing Technology

Bootstrap Confidence Interval of the Difference Between Two Process Capability Indices

J.-P. Chen[1] and L.-I. Tong[2]

[1]Department of Industrial Engineering and Management, National Chin-Yi Institute of Technology, Taiwan; and [2]Department of Industrial Engineering and Management, National Chiao Tung University, Taiwan

The process capability index, C_{pk}, is extensively used to assess process performance in manufacturing industry. Statistical hypothesis testing or an interval estimation of a single C_{pk} had been derived under a normally distributed process. However, the difference between two process capability indices, $C_{pk1} - C_{pk2}$, cannot be inferred statistically because of the complexity of the sampling probability theory. This study proposes a bootstrap resampling simulation method to construct the biased corrected percentile bootstrap (BCPB) confidence interval of ($C_{pk1} - C_{pk2}$), which can be used to select the better of two suppliers. The various simulation results indicate that the bootstrap confidence interval of ($C_{pk1} - C_{pk2}$) can be employed to evaluate effectively the difference between the performance of two suppliers.

Keywords: Bootstrap method; Interval estimation; Process capability index; Sensitivity analysis

1. Introduction

Internationalisation has become the goal of many companies and industries. Manufacturers must respond to the consumers' demands and offer after-sales service, as well as maintain the high quality and low cost of the product. Therefore, effective total quality management and accurate evaluation of manufacturers' process capabilities have become important in industry.

Process capability analysis is used to determine whether the process capability of a supplier conforms to a customer's requirement, by applying an expression called the process capability index (PCI), to a controlled process. Accordingly, the PCI can be used by both producer and supplier as a reference when signing a contract. Purchasing personnel can use the PCI to decide whether to accept or reject products provided by suppliers. Quality engineers can use the PCI to evaluate and continually improve a process.

Several process capability indices have been suggested to assess processes. Among them, C_p and C_{pk} [1,2] are frequently employed to evaluate process capability in manufacturing industries. However, C_{pk} is more extensively used in practice than is C_p because the former considers the degree of process mean, μ, shifted from the centre of the specifications and the standard deviation, σ, simultaneously. If the process is normally distributed, C_{pk} can be defined as follows:

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

where USL is the upper specification limit and LSL is the lower.

A larger PCI value implies a better process capability. For economic reasons, 100% inspection is impossible for most manufacturers, and hence the process mean, μ, and the standard deviation, σ, are unknown. The sample mean ($\overline{X}$) and sample standard deviation (S) are the unbiased estimators of μ and σ under the normal assumption, but the unbiased estimator of C_{pk} is difficult to obtain. The estimator of C_{pk} is defined as follows:

$$\hat{C}_{pk} = \min\left\{\frac{USL - \overline{X}}{3S}, \frac{\overline{X} - LSL}{3S}\right\} \tag{2}$$

Many methods have been developed methods for evaluating whether a single supplier's process conforms to a customer's requirements. However, very few studies have addressed the difference between two suppliers' PCIs. Two common methods are considered to obtain the difference between two suppliers' PCIs. They are described as follows:

1. 100% inspection is performed to calculate separately the PCI for each supplier; the suppliers are then compared according to their respective true PCI values.
2. A sample is inspected, and statistical testing is used to assess two suppliers' process capabilities.

Method 1 requires 100% inspection, which is very expensive and time-consuming, and is therefore usually not used in

Correspondence and offprint requests to: Dr. J.-P. Chen, Department of Industrial Engineering and Management, National Chin-Yi Institute of Technology, 35 Lane 215 Section 1, Chung Shan Road, Taiping, Taichung 411, Taiwan. E-mail: chen504@chinyi.ncit.edu.tw

practice. Method 2 is also difficult to implement since a proper statistical testing procedure may not exist for all PCIs. Chou [3] offered a procedure for comparing pairs of C_p, C_{pu} or C_{pl} indices. However, a procedure to compare two C_{pk} indices has not yet been developed, because the sampling distribution of the estimator of $C_{pk1} - C_{pk2}$ is more complex than that of the estimators of C_p, C_{pu}, and C_{pl}. Consequently, an exact statistical hypothesis test for $C_{pk1} - C_{pk2}$ cannot be constructed, indicating that method 2 is hard to implement for determining the difference between two C_{pk} indices. However, the confidence interval can then be used to assess the two suppliers' capabilities if the confidence interval of the difference between two suppliers' process capabilities can be obtained.

This study aims to obtain the confidence interval between the C_{pk} of two suppliers, using the bootstrap resampling simulation method [4]. The coverage proportion and other criteria are used to determine the accuracy of the method.

2. Schema of the Bootstrap Confidence Interval

Efron [5,6] introduced and developed a non-parametric, but computationally intensive, estimation method called "bootstrap". It is a databased simulation method for statistical inference. In particular, the non-parametric bootstrap can be used to estimate the sampling distribution of a statistic, while assuming only that the sample is representative of the population from which it is drawn and that the observations are independent and identically distributed. In its simplest form, the non-parametric bootstrap does not rely on any distributional assumptions regarding the underlying population. Let $\{x_1,x_2,\ldots,x_n\}$ be a sample of size n, taken from a process. A bootstrap sample, denoted by $\{x_1^*,x_2^*,\ldots,x_n^*\}$, is a sample of size n drawn (with replacement) from the original sample. Hence, a total of n^n resamples is possible.

Bootstrap sampling is equivalent to sampling (with replacement) from the empirical probability distribution function. Accordingly, the bootstrap distribution resembles the underlying distribution. In practice, only a random sample of the n^n possible resamples is commonly drawn. The estimates are determined for each process index, and the subsequent empirical distribution is referred to as the statistic's bootstrap distribution. Efron and Tibshirani [4] claimed that a minimum of approximately 1000 bootstrap resamples is usually sufficient for obtaining reasonably accurate confidence interval estimates.

Suppose we have a random variable, X, to evaluate the performance of a process. Although the distribution of X is unknown, we wish to estimate some parameter, θ, that characterises the performance of the process. An estimate of $\hat{\theta}$ can be determined using the bootstrap sample. The estimate is represented by $\hat{\theta}^*$, and is called the bootstrap estimate. The resampling procedure can be repeated many times, say, B times. The B bootstrap estimates $\hat{\theta}_1^*,\hat{\theta}_2^*,\ldots,\hat{\theta}_B^*$ can be calculated from the resamples. Other studies on bootstrap methods include Efron and Gong [7], Gunter [8,9], Mooney and Duval [10], or Young [11].

Efron and Tibshirani [4] further developed three types of bootstrap confidence interval – the standard bootstrap (SB) confidence interval, the percentile bootstrap (PB) confidence interval, and the biased corrected percentile bootstrap (BCBP) confidence interval. The relevant formulae are as follows:

1. Standard Bootstrap (SB). The sample mean and the standard deviation of B bootstrap estimates $\hat{\theta}_i^*$ can be obtained as follows.

$$\hat{\theta}^* = \frac{1}{B}\sum_{i=1}^{B}\hat{\theta}_i^*,$$

$$S_{\hat{\theta}}^* = \sqrt{\frac{1}{B-1}\sum_{i=1}^{B}(\hat{\theta}_i^* - \hat{\theta}^*)^2}$$

If the distribution of $\hat{\theta}$ is approximately normal, the $(1-\alpha)$ 100% SB confidence interval for θ is $\hat{\theta}^* \pm z_{\frac{\alpha}{2}}S_{\hat{\theta}}^*$ where $z_{\frac{\alpha}{2}}$ is the 100 $(\alpha/2)^{\text{th}}$ percentage point of the standard normal distribution.

2. Percentile Bootstrap (PB). From the ordered set of $\hat{\theta}_1^*$, the $(1-\alpha)$ 100% PB confidence interval for θ can be obtained as follows:

$$[\theta^*(\alpha/2 \times B), \hat{\theta}^*((1-\alpha/2) \times B)]$$

where $\hat{\theta}^*(i)$ is the ith value of ordered $\hat{\theta}_i^*$, $i = 1, 2,\ldots, B$.

3. Biased-Corrected Percentile Bootstrap (BCPB). The bootstrap distribution determined from only a sample of the complete bootstrap distribution may be shifted higher or lower than expected That is, the distribution is biased. Accordingly, a third method has been presented to correct this potential bias (see [6] for a complete justification of the method.) First, the distribution of $\hat{\theta}^*(i)$ is used to calculate the probability,

$$p_0 = Pr[\hat{\theta}^*(i) \le \hat{\theta}] \quad (i = 1,2,\ldots,B)$$

where $\hat{\theta}$ is the value of θ estimated from a random sample $\{x_1,x_2,\ldots,x_n\}$. Secondly, the following are calculated:

$$z_0 = \Phi^{-1}(p_0)$$
$$P_L = \Phi(2z_0 - z_{\alpha/2})$$
$$P_U = \Phi(2z_0 + z_{\alpha/2})$$

where $\Phi(\bullet)$ is the cumulative standard normal distribution function. Finally, the BCPB confidence interval is obtained as follows.

$$[\hat{\theta}^*(P_L B), \hat{\theta}^*(P_U B)]$$

Chou [3] proposed three one-sided tests of the difference between pairs of C_p, C_{pu}, and C_{pl} process capability indices under the normal assumptions. For testing the null hypothesis H_0: $C_{p_1} \ge C_{p_2}$ against the alternative hypothesis H_a: $C_{p1} < C_{p_2}$, the test statistic is given by $F_0 = S_1^2/S_2^2$, which has an F-distribution with $(n-1)$ and $(m-1)$ degrees of freedom when $\sigma_1 = \sigma_2$. For a given significance level, α, H_0 is rejected if $F_0 > F_{(1-\alpha,n-1,m-1)}$, where $F_{(1-\alpha,n-1,m-1)}$ is the $100(1-\alpha)$th percentage point of the F-distribution with $(n-1)$ and $(m-1)$ degrees of freedom. In this case, supplier 2 is more capable than supplier 1. Furthermore, to test the null hypothesis H_0: $C_{pu_1} \ge C_{pu_2}$ (or $C_{pl_1} \ge C_{pl_2}$) against the alternative hypothesis H_a: $C_{pu_1} < C_{pu_2}$ (or $C_{pl_1} < C_{pl_2}$), the sampling distribution of the test statistic is very complex. The sampling distributions of $\hat{C}_{pk}$ or $(\hat{C}_{pk1} - C_{pk2})$ are much more complicated than those

Table 1. The simulation results of a BCPB confidence interval for various process parameter combinations.

μ_1	μ_2	σ_1	σ_2	Sample sizes	Coverage proportion	Average length of the interval width	Standard deviation of the interval width
310	305	10	9	20	0.942	1.34142	0.31028
				40	0.937	0.86816	0.13436
				60	0.937	0.68159	0.08416
		12	7	20	0.922*	1.39725	0.31630
				40	0.934	0.91020	0.14723
				60	0.944	0.72381	0.09511
		14	5	20	0.913*	1.64495	0.39445
				40	0.932*	1.07735	0.19168
				60	0.935	0.87772	0.13125
315	300	10	9	20	0.935	1.26984	0.29246
				40	0.937	0.81212	0.12501
				60	0.955	0.64506	0.08401
		12	7	20	0.93*	1.28938	0.30407
				40	0.929*	0.82535	0.12802
				60	0.933	0.65382	0.07938
		14	5	20	0.908*	1.44659	0.35180
				40	0.926*	0.95289	0.16090
				60	0.943	0.76534	0.10817
320	295	10	9	20	0.932*	1.11397	0.24152
				40	0.937	0.71538	0.11092
				60	0.941	0.56659	0.06965
		12	7	20	0.928*	1.13815	0.24420
				40	0.937	1.12910	0.26129
				60	0.953	0.57602	0.07199
		14	5	20	0.932*	1.29113	0.31908
				40	0.935	0.82741	0.13414
				60	0.94	0.66377	0.08792

*Indicates a proportion significantly different from the expected value of 0.950 (at $\alpha = 0.01$).

of C_{pu} and C_{pl}. No study has yet proposed either a method for developing the unbiased estimators of $\hat{C}_{pk1} - C_{pk2}$, or the sampling distributions of test statistics.

Franklin and Wasserman [12] offered three non-parametric bootstrap lower confidence limits for C_p and C_{pk}. A simulation was performed, and the bootstrap and the parametric estimates were compared. The simulation results revealed that, in the normal process environment, the bootstrap confidence limits perform as well as the lower confidence limits derived by the parametric method and based on a normal process (see Chou et al [13] for C_p; Bissell [14] for C_{pk}; Boyles [15] for C_{pm}). In non-normal process environments, the bootstrap estimates were significantly better.

In sum, the difference between two suppliers' C_{pk} cannot be evaluated by employing the procedure for testing the hypothesis H_0: $C_{pk1} \geqq C_{pk2}$ against H_a: $C_{pk1} < C_{pk2}$. If both the lower and upper confidence limits for the difference between two process capability indices ($C_{pk1} - C_{pk2}$), are positive, then supplier 1 has a better process capability than supplier 2; if both confidence limits are negative, then supplier 2 has a better process capability than supplier 1; if one confidence limit is positive and the other negative, then no significant difference exists between the two suppliers' process capabilities. In this study, the biased-corrected percentile bootstrap (BCP) is employed to determine the confidence limits for ($C_{pk1} - C_{pk2}$) and the results are used to select the better of the two candidates.

3. Procedures for Selecting a Better Supplier Using the Bootstrap Confidence Interval

This section presents a procedure for constructing a bootstrap confidence interval to select a better supplier using the bootstrap confidence limits. The accuracy and sensitivity of the procedure are analysed for various sample sizes, process means and standard deviations.

3.1 BCPB Confidence Interval for $C_{pk1} - C_{pk2}$

If two suppliers' processes are independently normally distributed and the target value of each process is the specification centre, then the bootstrap confidence interval for the difference between two suppliers' process capability indices, $C_{pk1} - C_{pk2}$, can be obtained as follows:

Step 1. Assume that the two suppliers' processes are normally distributed as $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$. The target value is T, and the specification centre is m ($m = (USL - LSL)/2$).

Step 2. Randomly select m_1 and m_2 from suppliers 1 and 2, respectively. These are referred to as the original samples.

Step 3. Select the bootstrap sample with a replacement from each original sample, by the bootstrap resampling method.

Step 4. Compute the jth ($j = 1,\ldots,B$) estimates of C_{pk1} and C_{pk2} from the jth bootstrap sample. These are defined as $\hat{C}^*_{pk1j}$ and $\hat{C}^*_{pk2j}$. The difference, $\hat{e}^*_j = \hat{C}^*_{pk1j} - \hat{C}^*_{pk2j}$ is also calculated.

Step 5. Repeat Steps 3 and 4 B times ($B = 1000$). The B bootstrap samples are obtained and the set of B values of $\hat{e}^*_j$ are sorted in increasing order, and referred to as $\hat{e}^*(1)$,$e^*(2)$,$\ldots$,$\hat{e}^*(B)$.

Step 6. Find the $(1 - \alpha)$ 100% BCPB confidence interval of ($C_{pk1} - C_{pk2}$), that is $[\hat{e}^*(P_L B), \hat{e}^*(P_U B)]$.

Statistical analysis software, SAS, is used to generate the random samples from two independent normal processes and thus calculate the BCPB confidence interval for ($C_{pk1} - C_{pk2}$). An example of the computation follows.

Step 1. Assume that two suppliers' process distributions are from $N_1(315, 12)$ and $N_2(300, 7)$. The upper and lower specification limits are 353 and 273, and the specification centre is 313, which is also the target value of the process.

Step 2. Generate two original random samples of size 40 from $N_1(315, 12)$ and $N_2(300, 7)$, respectively: $\{x_1, x_2, \ldots, x_{40}\}_1$ and $\{x_1, x_2, \ldots, x_{40}\}_2$.

Step 3. Generate bootstrap samples $\{x^*_1, x^*_2, \ldots, x^*_{40}\}_1$, $\{x^*_1, x^*_2, \ldots, x^*_{40}\}_2$ by bootstrap resampling with replacement from the original random samples, $\{x_1, x_2, \ldots, x_{40}\}_1$ and $\{x_1, x_2, \ldots, x_{40}\}_2$.

Step 4. Compute the difference $\hat{e}^* = \hat{C}^*_{pk1} - \hat{C}^*_{pk2}$.

Step 5. Repeat steps 3 and 4 one thousand ($B = 1000$) times. The estimates, $\hat{e}^*_j$, $j = 1,\ldots,1000$, are sorted in ascending order, $\hat{e}^*(1),\ldots,\hat{e}^*(1000)$. Table 1 gives the results.

Step 6. Obtain a 95% BCPB confidence limits for ($C_{pk1} - C_{pk2}$) as (-0.33381, 0.5563).

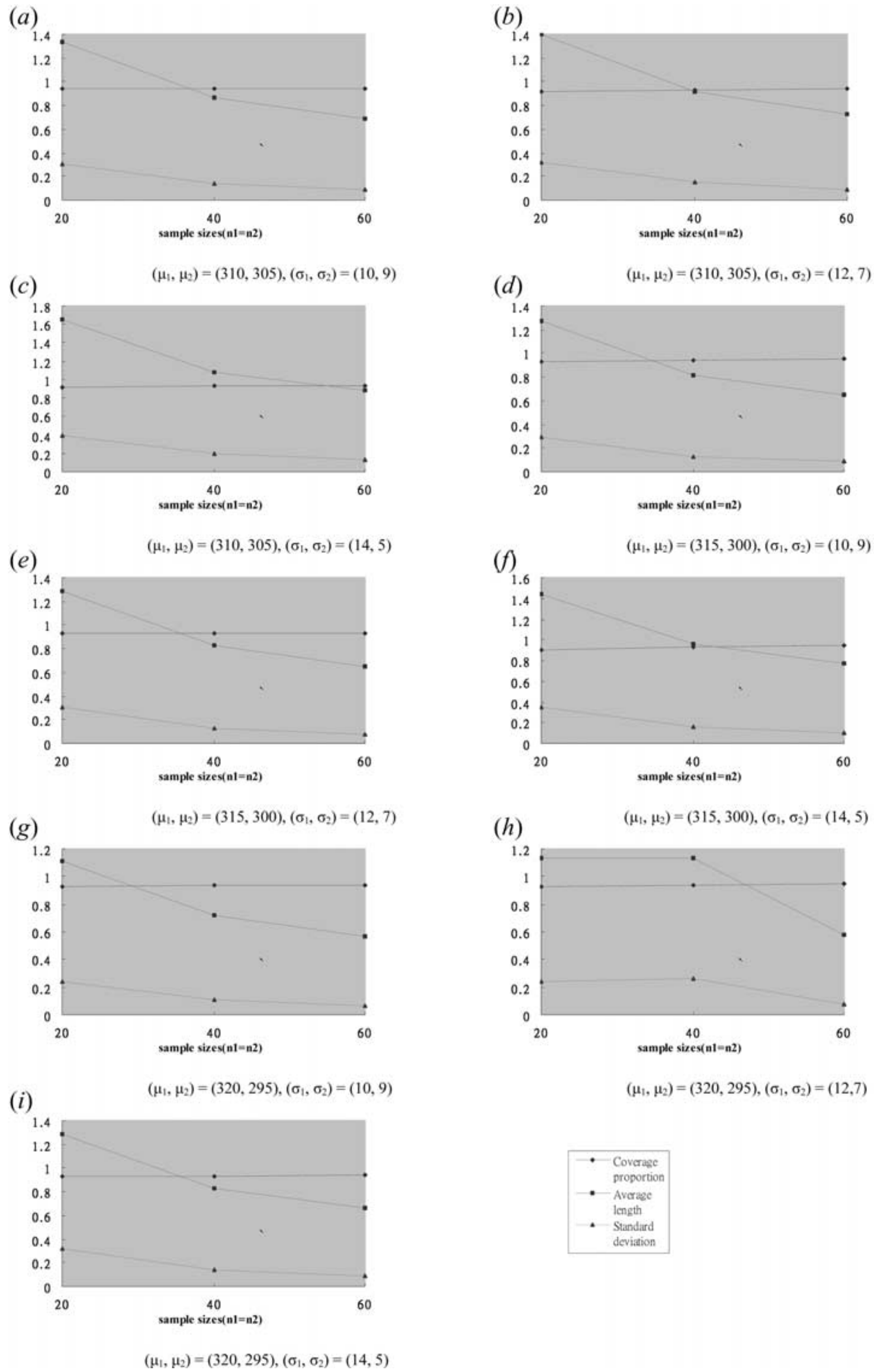

Fig. 1. The trend diagrams of three criteria for various sample sizes.

3.2 Analysing the Accuracy and Sensitivity of the BCPB Confidence Interval

Monte Carlo simulation is used to determine the accuracy of the BCPB confidence interval for $(C_{pk1} - C_{pk2})$, as follows:

1. Repeat steps 3 to 6 in Section 3.1, N times (a larger N yields greater effectiveness). N sets of BCPB confidence intervals are obtained.
2. The actual difference between two suppliers' process capability indices, $C_{pk1} - C_{pk2}$, is computed as follows:

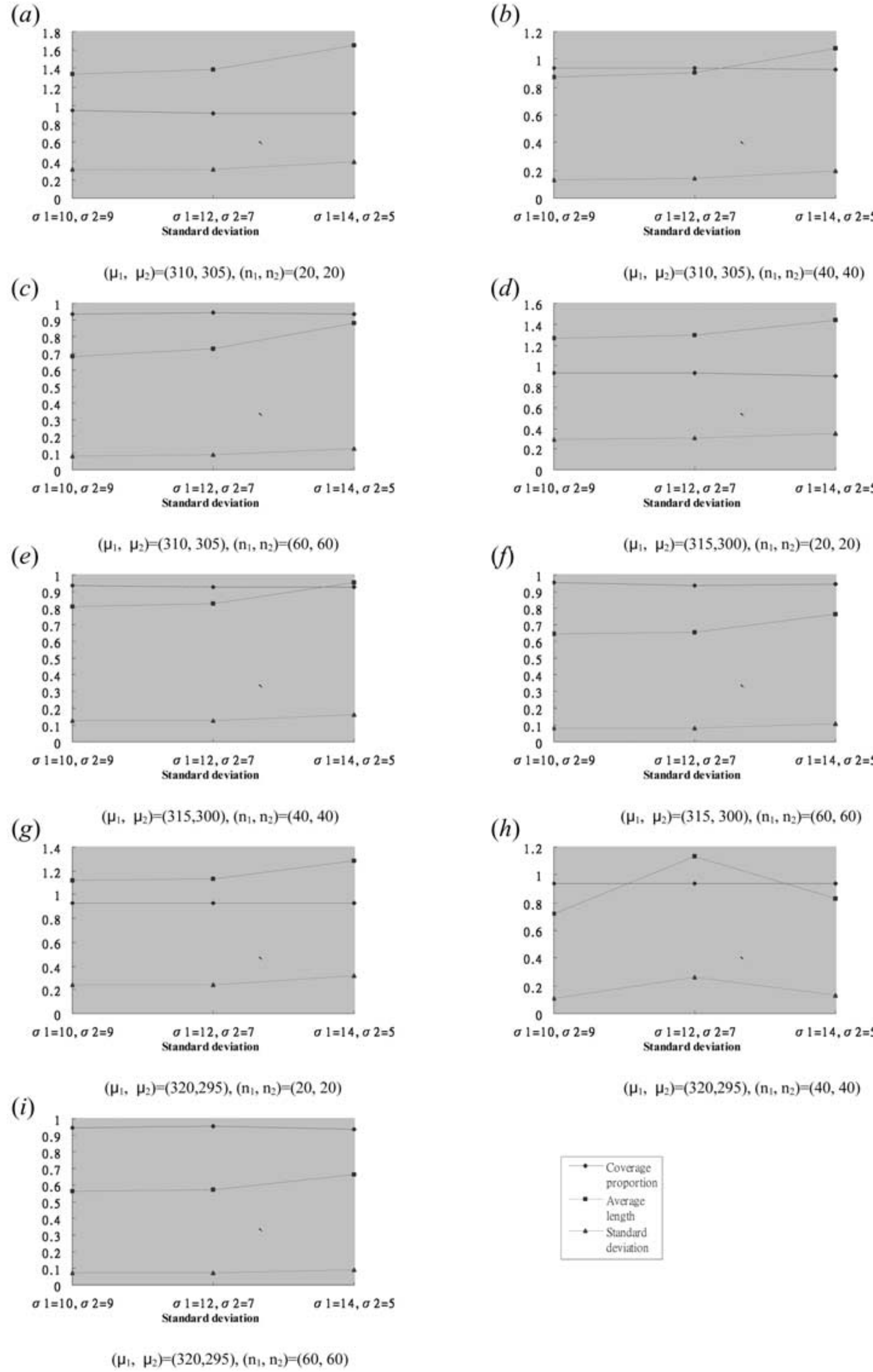

Fig. 2. The trend diagrams of three criteria for various standard deviations.

$$D = \min\left\{\frac{USL - \mu_1}{3\sigma_1}, \frac{\mu_1 - LSL}{3\sigma_1}\right\} - \min\left\{\frac{USL - \mu_2}{3\sigma_2}, \frac{\mu_2 - LSL}{3\sigma_2}\right\}$$

3. The actual proportion of the $(1 - \alpha)100\%$ BCPB confidence intervals that contain D are computed.
4. The actual proportion is compared to $(1 - \alpha)100\%$.

If the actual proportion is greater than or equal to $(1 - \alpha)100\%$, then the BCPB confidence interval method is considered effective and accurate.

The simulation method is also used to analyse the sensitivity of the proposed BCPB confidence interval. Sensitivity analysis can be performed by repeating steps 3 to 6 for various para-

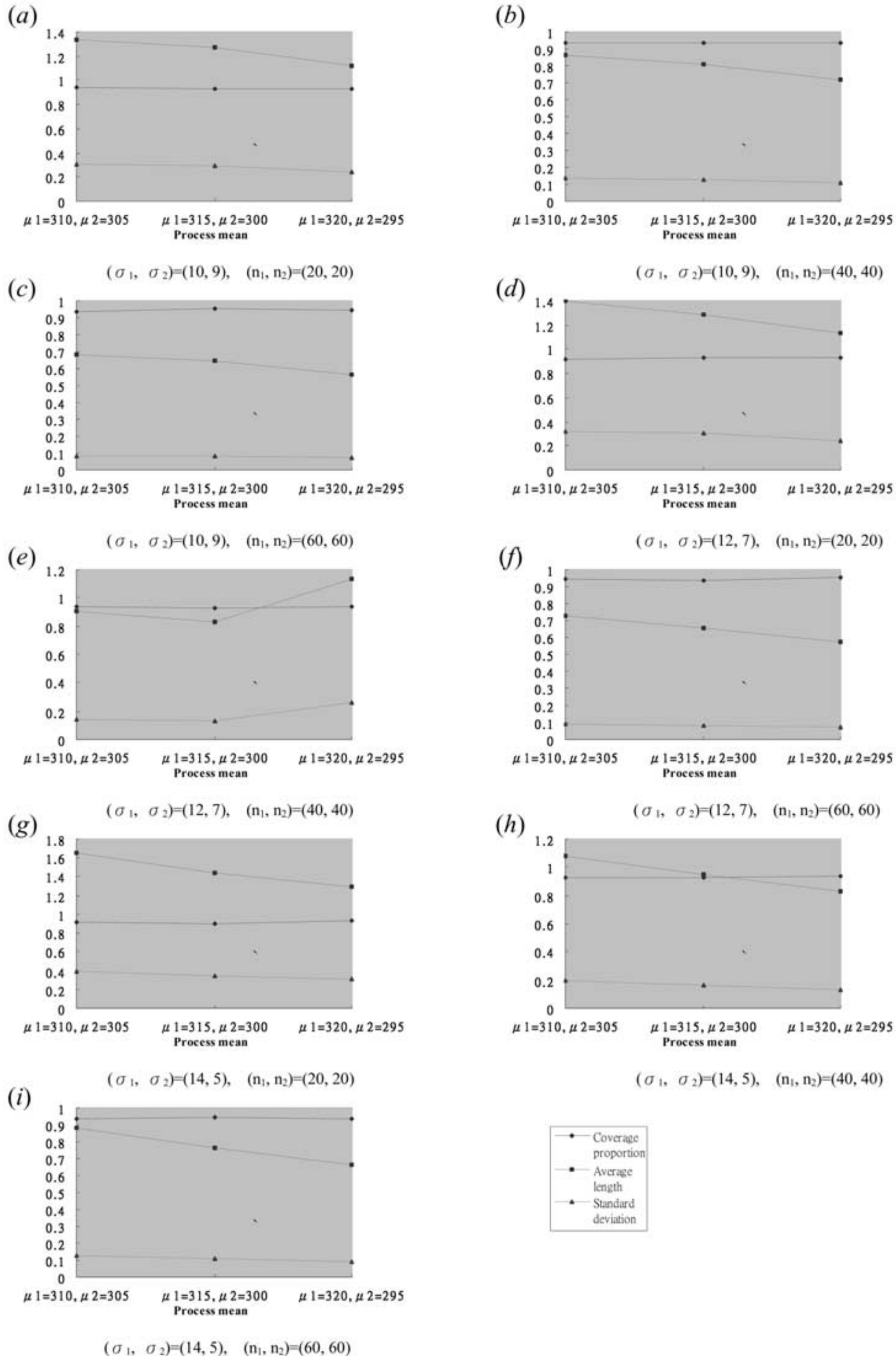

Fig. 3. The trend diagrams of three criteria for various process means.

meter combinations of sample size, process mean, and process standard deviation. The accuracy and effectiveness of the BCPB confidence interval are expected to be robust for various process parameter combinations.

The accuracy and sensitivity of various process parameters $(\mu_1, \mu_2) = (310, 305)(315, 300)(320, 295)$, $(\sigma_1, \sigma_2) = (10, 9)(12, 7)(14, 5)$, $(n_1, n_2) = (20, 20)(40, 40)(60, 60)$ are considered.

Table 1 displays the simulation results for a 95% BCPB confidence interval under various parameter combinations. Three criteria are used to measure the performance of the BCPB confidence interval:

1. Coverage proportion. The coverage proportion is the number of times that the BCPB confidence intervals contained the

(*a*) (*b*)

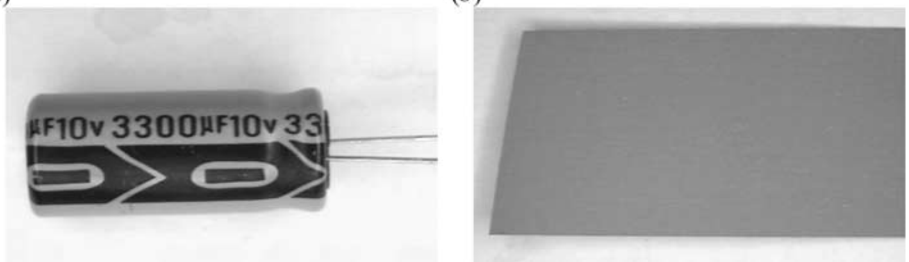

Fig. 4. (*a*) The capacitor and (*b*) aluminium foil.

Table 2. Data of voltages for aluminium foils from two suppliers. (USL, T, LSL) = (530, 520, 510), (n_1, n_2) = (50, 50), measurement unit: WV.

Supplier 1					Supplier 2				
519.9	519.5	520.1	517.0	521.1	521.7	521.3	523.5	524.4	522.5
517.1	518.7	520.1	521.2	521.7	523.3	527.1	524.9	522.9	524.2
520.4	517.9	522.9	517.7	517.2	523.9	523.5	527.5	517.3	518.7
520.7	521.0	519.1	518.4	518.9	518.7	521.9	519.7	520.4	520.4
517.9	518.4	520.8	519.3	520.6	523.7	526.8	517.7	528.1	517.5
516.6	519.0	520.6	517.9	519.6	523.8	514.7	522.6	518.5	526.3
519.6	522.6	518.3	522.1	523.1	523.2	524.4	522.7	519.6	520.4
519.9	519.8	520.7	516.5	521.5	520.6	525.2	524.1	519.3	522.2
519.2	521.2	518.9	517.8	521.3	520.1	521.9	516.7	520.9	525.2
521.3	517.4	519.5	522.0	523.8	522.6	523.1	521.7	520.9	526.3

true difference value, D, over $N = 1000$ runs. A larger coverage proportion corresponds to a better performance.

2. Average length. The average length of the BCPB confidence interval is calculated for $N = 1000$ runs. A shorter average length corresponds to a better performance.
3. Standard deviation. The standard deviation of the lengths of the bootstrap confidence interval is calculated for $N = 1000$ runs. A lower standard deviation corresponds to a better performance.

The frequency of coverage for the confidence interval is binomially distributed with $N = 1000$ and $p = 0.95$. Hence, a 99% confidence interval for the coverage rate is $0.95 \pm 2.575 \times \sqrt{((0.95 \times 0.05)/1000)} = 0.95 \pm 0.0178$, Hence, we could be 99% confident that a "true 95% confidence interval" would have a coverage proportion between 0.933 and 0.967. Although ten out of 27 possible intervals are below 0.933, they all exceed 0.91, and the simulation results are considered reliable.

The results of the sensitivity analysis are as follows:

1. Figures 1(*a*) to 1(*i*) display the trend of three criteria for various sample sizes with given pairs of process mean and standard deviation values.

   Figures 1(*a*) to 1(*i*) reveal that the coverage proportion increases slightly, and the average length and standard deviation of the intervals decreases, as the sample size increases. The accuracy of the BCPB confidence intervals also increases with sample size. For $n = 20$, 1 out of the 9 computed BCPB confidence limits contains the actual D value. For $n = 40$, 4 out of the 9 computed BCPB confidence limits contains the D value. For $n = 60$, 7 out of 9 computed BCPB confidence limits contains the D value. Accordingly, the sample size for the BCPB confidence interval of ($C_{pk1} - C_{pk2}$) should be greater than or equal to 40.
2. Figures 2(*a*) to (*i*) display the trend of three criteria for various combinations of standard deviations with particular sets of process means and various sample sizes.

   Figures 2(*a*) to 2(*i*) reveal that the coverage proportion decreases as σ_1/σ_2 increases. Conversely, the average length and standard deviation increase as σ_1/σ_2 increases, except in the case where (μ_A, μ_B) = (320, 295) and (n_1,n_2) = (40,40).
3. Figures 3(*a*) to 3(*i*) display the trend of three criteria for various combinations of process means with particular sets of process standard deviations and sample sizes.

Figures 3(*a*) to (*i*) reveal that the coverage proportion does not change significantly as (μ_1,μ_2) increases. Generally, the average length and standard deviation decreases as (μ_1,μ_2) increases, except in cases where (σ_A, σ_B) = (12, 7) and (n_1,n_2) = (40, 40). With respect to the accuracy of BCPB confidence intervals, although 10 out of the 27 coverage proportions are outside the true 95% BCPB confidence intervals, the coverage proportions for various process parameter combinations all exceed 0.91, indicating that the BCPB confidence interval can be used as a reliable tool for selecting the more capable of two competing suppliers.

In the sensitivity analysis of the BCPB confidence intervals, the simulation results reveal that the behaviour of the three evaluating criteria (coverage proportion, average length, and standard deviation) changes very little for various parameter combinations of process mean, process standard deviation, and sample sizes. The findings show that the BCPB confidence interval is robust. We recommend that sample sizes above 40 are necessary for employing the BCPB confidence interval for $(C_{pk1} - C_{pk2})$.

4. Numerical Example

This study cites data from two suppliers, who provided aluminium foil materials to an electronics company in Taiwan, to demonstrate the proposed procedure. Aluminium foil is a key component that governs the quality of capacitors (see Fig. 4), and the voltage is an important quality characteristic of aluminium foil: the production specifications (USL, T, LSL) of the voltage are (530, 520, 510). If the voltage falls outside this interval, the aluminium foil will break, and thus be rejected. Fifty random samples are taken from suppliers 1 and 2 by a quality inspector. Table 2 shows the collected sample data. The process of each supplier is approximately normally distributed. The simulation result for the 95% BCPB confidence interval for $C_{pk1} - C_{pk2}$ is (0.44139, 1.18139). We are thus 95% confident that supplier 1 is more capable than supplier 2, since the lower and upper confidence limits are positive.

5. Conclusions

The process capability indices, C_p and C_{pk}, are extensively used in manufacturing industries. C_{pk} simultaneously measures the ability of a process to meet a required target value, and the variation within specified limits. Therefore, C_{pk} more accurately assesses the process capability than C_p. Although statistical tests have been developed to compare two C_p, C_{pu} and C_{pl} process capability indices of normal processes, a statistical test for comparing two C_{pk} values has not been developed due to the complexity of the sampling distribution of $(\hat{C}_{pk1} - \hat{C}_{pk2})$.

This study proposes a BCPB confidence interval for $(C_{pk1} - C_{pk2})$ that involves the bootstrap resampling method to replace the hypothesis testing method for selecting the more capable of two candidate suppliers. This bootstrap interval is accurate and effective when used for this purpose. Quality engineers or managers with limited statistical background can easily implement the procedure for establishing the BCPB confidence interval and selecting a capable supplier.

References

1. H. C. Charbonneau and G. L. Webster, Industrial Quality Control, Prentice-Hall, Englewood Cliffs, New Jersey, 1978.
2. V. E. Kane, "Process capability indices", Journal of Quality Technology, 18(1), pp. 41–52, 1986.
3. Y. M. Chou, "Selecting a better supplier by testing process capability indices", Quality Engineering, 6(3), pp. 427–438, 1994.
4. B. Efron and R. J. Tibshirani, "Bootstrap methods for standard errors, confidence interval, and other measures of statistical accuracy", Statistical Science, 1, pp. 54–77, 1986.
5. B. Efron, "Bootstrap methods: another look at the jackknife", Annals of Statistics, 7, pp. 1–26, 1979.
6. B. Efron, The Jacknife, the Bootstrap and Other Resampling Plans, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982.
7. B. Efron and G. Gong, "A leisurely look at the bootstrap, the jack-knife, and cross-validation", American Statistician, 37, pp. 36–48, 1983.
8. B. Gunter, "Bootstrapping: how to make something from almost nothing and get statistically valid answers, Part I", Quality Progress, pp. 97–103, 1991.
9. B. Gunter, "Bootstrapping: how to make something from almost nothing and get statistically valid answers, Part III", Quality Progress, pp. 119–122, 1992.
10. C. Z. Mooney and R. D. Duval, Bootstrapping: A Nonparametric Approach to Statistical Inference, SAGE, Newbury Park, CA, 1993.
11. G. A. Young, "Bootstrap: more than a stab in the dark?", Statistical Science, 9, pp. 382–415, 1994.
12. L. A. Franklin and G. S. Wasserman, "Bootstrap lower confidence limits for capability indices", Journal of Quality Technology, 24(4), pp. 196–210, 1992.
13. Y. Chou, D. B. Owen and S. A. Borrego, "Lower confidence limits on process capability indices", Journal of Quality Technology, 22, pp. 223–229, 1990.
14. A. F. Bissell, "How reliable is your capability index?", Applied Statistics, 39, pp. 331–340, 1990.
15. R. A. Boyles, "The Taguchi capability index", Journal of Quality Technology, 23, pp. 17–26, 1991.