QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL, VOL. 13, 337–346 (1997)

# A GRAPHICAL APPROACH TO OBTAINING CONFIDENCE LIMITS OF $C_{pk}$

loon ching tang, su ee than and beng wah ang
*Department of Industrial and Systems Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260*

SUMMARY

The process capability index $C_{pk}$ has been widely used as a process performance measure. In practice this index is estimated using sample data. Hence it is of great interest to obtain confidence limits for the actual index given a sample estimate. In this paper we depict graphically the relationship between process potential index ($C_p$), process shift index ($k$) and percentage non-conforming ($p$). Based on the monotone properties of the relationship, we derive two-sided confidence limits for $k$ and $C_{pk}$ under two different scenarios. These two limits are combined using the Bonferroni inequality to generate a third type of confidence limit. The performance of these limits of $C_{pk}$ in terms of their coverage probability and average width is evaluated by simulation. The most suitable type of confidence limit for each specific range of $k$ is then determined. The usage of these confidence limits is illustrated via examples. Finally a performance comparison is done between the proposed confidence limits and three non-parametric bootstrap confidence limits. The results show that the proposed method consistently gives the smallest width and yet provides the intended coverage probability. © 1997 John Wiley & Sons, Ltd.

key words: process capability analysis: $C_{pk}$; confidence interval; six sigma

## INTRODUCTION

Process capability indices $C_p$ and $C_{pk}$ are commonly used to determine whether a process is capable of producing within specification limits. Capability indices can be used to relate the process parameters $\mu$ and $\sigma$ to engineering specifications that may include unilateral or bilateral tolerances with or without a target (or nominal) value. Kane[1] noted that these indices provide an easily understood language for quantifying the performance of a process. The capability measurement is compared with the tolerance in order to judge the adequacy of the process. Sullivan[2] proposed that a process with high process capability index values would effectively eliminate inspection and defective material and therefore eliminate costs associated with inspection, scrap and rework.

Montgomery[3] presented some recommended guidelines for minimum values of $C_p$. If the estimated indices are larger than or equal to the respective recommended minimum values, then we claim that the process is capable. For instance, if one desires a non-conforming (NC) proportion of 0·007 per cent, then a minimum value of $C_p = 1{\cdot}33$ is recommended for an ongoing process, assuming that the process is entered. However, owing to sampling error, the estimated indices $\hat{C}_p$ and $\hat{C}_{pk}$ are likely to be different from the true indices $C_p$ and $C_{pk}$. The recommended minimum values are for the true indices. Hence, even when the estimated index is larger than the minimum value, one may not be able to claim with a high level of confidence that the true index is indeed larger than the minimum value. This problem arises because the sampling error increases when the sample size used for estimation decreases. Thus a mere comparison between the estimated index and the recommended minimum values may not be a good process capability gauge, particularly when the sample size is small. This situation may arise in a short-run production where the initial samples available for process qualification are limited.

The above problem can be overcome by using lower confidence limits for $C_{pk}$. Unfortunately, it is not straightforward to obtain a confidence bound for $C_{pk}$. Difficulties in the construction of confidence limits for $C_{pk}$ arise from the rather complicated way in which $\mu$ and $\sigma$ appear in the expression for $C_{pk}$. Wang[4] pointed out that this complication has made the development of confidence limits for $C_{pk}$ a much more difficult problem than simply combining the confidence limits of $\mu$ and $\sigma$ to give a maximum and minimum $C_{pk}$.

Several methods for constructing lower confidence limits for $C_{pk}$ have been presented in the literature. Reasonably comprehensive comparisons of these methods are given by Kushler and Hurley.[5] In that study, comparisons of the Chou approach,[6] the non-central $t$-distribution approach and other approaches were made by numerical integration of the joint density function of a sample mean and a sample

CCC 0748–8017/97/060337–10 $17.50

*Received 17 January 1997*
*Revised 30 April 1997*

variance. The methods were then compared by their miss rates, i.e. the frequency with which the $1-\alpha$ confidence limits constructed by each method exclude true $C_{pk}$-values. Kushler and Hurley concluded that the Chou approach is overly conservative. The non-central $t$-distribution method, which is exact for unilateral specification, gives a better approximation when the process is operating appreciably off-centre.

Three normal approximations to the sampling distribution of $C_{pk}$ are available in the literature. One involving gamma functions was derived by Zhang *et al*.[7] An alternative expression developed by Bissell[8] produces remarkably accurate lower $100(1-\alpha)\%$ confidence limits for $C_{pk}$ for $n \geq 30$. The formula derived is relatively easy to compute and is based on a Taylor series. The normal approximation presented by Kushler and Hurley[5] is a simplification of Bissell's result.

Another approach, also by Kushler and Hurley,[5] used the fact that since $C_{pk}$ depends simultaneously on $\mu$ and $\sigma$, a joint confidence region for these two parameters can be used to obtain a confidence bound for $C_{pk}$. The minimum value of $C_{pk}$ over the region is the lower confidence bound. One standard way to define a joint confidence region is to use contours of the likelihood function from the likelihood ratio test. Such confidence regions are asymptotically optimal and have also been found to provide good small-sample behaviour in a variety of problems. However, determining the confidence bound by this method requires complex calculations.

Franklin and Wasserman[9] proposed a non-parametric but computer-intensive bootstrap estimation to develop three types of bootstrap confidence intervals for $C_{pk}$. the standard bootstrap interval (SB), the percentile bootstrap confidence interval (PB) and the bias-corrected percentile bootstrap confidence interval (BCPB). They presented an initial study of the properties of these three bootstrap confidence intervals. The results indicated that some of the 90% bootstrap intervals provided 90% coverage when the process was normally distributed and provided 84%–87% when the process was chi-square distributed. Practitioners, however, have to bear in mind that the practical interpretation of the index $C_{pk}$ is questionable when normality does not hold.

To choose a method for determining confidence limits for process capability indices, Kushler and Hurley[5] considered ease of use and method performance as guidelines. An additional guideline which has so far been overlooked is that the method for constructing a confidence interval for $C_{pk}$ may depend on the nature of the process, e.g. a short-run production process or homogeneous batch process. Here we propose a scheme for determining the confidence interval that takes this guideline into consideration.

In the following section we depict graphically the relationship between $C_p$, $k$ and $p$. Based on the monotone property of the relationship, we derive the confidence limits of $k$ and construct a graphical tolerance box that relates the confidence limits of $C_p$, $k$ and $p$. Two different types of two-sided confidence limits for $C_{pk}$ are constructed by exploiting the fact that the sampling error in $C_{pk}$ originates independently of the sampling error in $k$ and $C_p$. These confidence limits are then combined using the Bonferroni inequality to give a conservative third type of two-sided confidence limit. The effectiveness of these confidence limits in terms of their coverage probability and average width is evaluated via simulations. By observing the performance of these confidence limits over different ranges of $k$, we propose an approximation method (AM) for choosing the appropriate confidence limits for $C_{pk}$. We also illustrate the usage of AM confidence limits via two examples. Finally we compare the performance of the proposed AM confidence limits with that of the bootstrap confidence limits developed by Franklin and Wasserman.[9] A discussion on the simulation results and recommendations on the type of confidence limits for different scenarios are presented.

## GRAPHING $C_p$, $k$ AND $p$

Consider a measured characteristic $X$ with lower and upper specification limits denoted by $LSL$ and $USL$ respectively. Measured values of $X$ that fall outside these limits will be termed 'non-conforming' (NC). An indirect measure of potential capability to meet the settings $LSL < X < USL$ is the process capability index

$$C_p = \frac{USL - LSL}{6\sigma} \tag{1}$$

Here $\sigma$ denotes the standard deviation of $X$ and is estimated by

$$s = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})^2} \tag{2}$$

where $n$ is the sample size and

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \tag{3}$$

Denoting the midpoint of the specification range by $m = (USL + LSL)/2$, the shift index $k$ is given by

$$k = \frac{|m - \mu|}{(USL - LSL)/2} \tag{4}$$

The proportion of NC product can be estimated in terms of $C_p$ and $k$ by

$$\hat{p} = \Phi\,[-3(1+\hat{k})\hat{C}_p] + \Phi\,[-3(1-\hat{k})\hat{C}_p] \tag{5}$$

A contour plot of (5) on two different scales

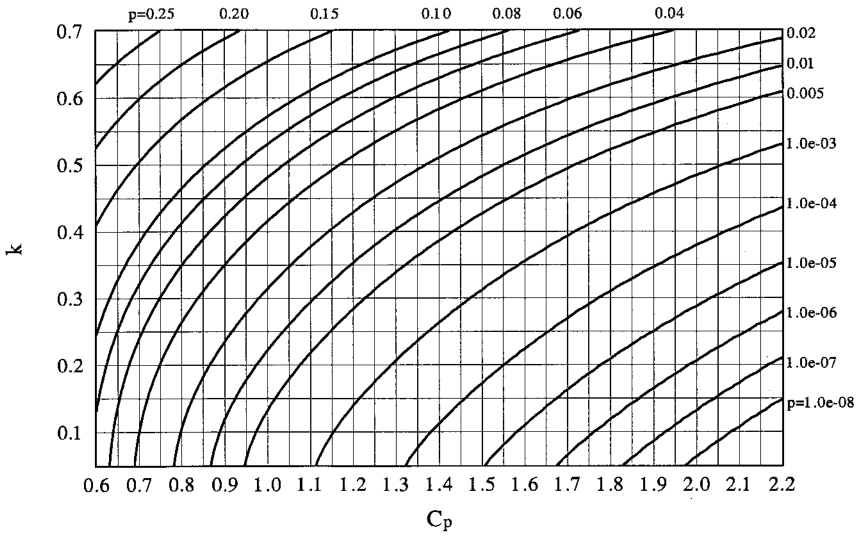


Figure 1. Behaviour of $C_p$, $k$ and $p$ on linear scale

(Figures 1 and 2) reveals some interesting and useful properties.

It can be seen from the figures that while $(C_p,k)$ uniquely determine $p$, there also exists a unique value of $k$ for each $(C_p,p)$. Intuitively, we know that for each constant $k$, $p$ increases as $C_p$ decreases, and for each constant $C_p$, $p$ increases as $k$ increases. Figure 1 also reveals that for each constant $p$, $k$ is monotonic increasing with respect to $C_p$. This unique relationship between $C_p$, $k$ and $p$ allows us to derive the confidence limits of $k$ from the easily computable confidence limits of $C_p$ and $p$.

## CONFIDENCE LIMITS FOR $k$

Consider a sample from a normally distributed stable process. A $100(1-\alpha)\%$ confidence interval for $C_p$ is

$$\left[\frac{USL-LSL}{6\hat{\sigma}}\frac{\chi^2_{n-1,\alpha/2}}{\sqrt{n-1}},\ \frac{USL-LSL}{6\hat{\sigma}}\frac{\chi^2_{n-1,1-\alpha/2}}{\sqrt{n-1}}\right]$$

$$\equiv\left[\frac{\chi^2_{n-1,\alpha/2}}{\sqrt{n-1}}\hat{C}_p,\ \frac{\chi^2_{n-1,1-\alpha/2}}{\sqrt{n-1}}\hat{C}_p\right] \qquad (6)$$

where $\chi^2_{n-1}$ is the chi-square distribution value with $n-1$ degrees of freedom. From the same sample the MLE of the fraction NC is given by $\hat{p}$, which is the number of NC units divided by $n$. The upper and lower confidence limits ($\bar{p}$ and $\underline{p}$) of $p$ can be obtained from the binomial equation.

Treating $k$ as the statistic of interest, we write

$$L(k:p,C_p)=p-\Phi\,[-3(1+k)C_p] + \Phi\,[-3(1-k)C_p] \qquad (7)$$

The unique and monotonic relationship between $k$ and $(C_p,p)$ permits the derivation of the upper and lower confidence limits of $k$, $(\bar{k},\underline{k})$, using $(\bar{C}_p,\bar{p})$ and $(\underline{C}_p,\underline{p})$ respectively. This is shown graphically in Figure 3 and can be expressed mathematically as

$$\underline{k}=\{k : L\,(k : \underline{p},\underline{C}_p)=0\} \qquad (8)$$

$$\bar{k}=\{k : L\,(k : \bar{p},\bar{C}_p)=0\} \qquad (9)$$

By choosing an appropriate significance level for the confidence limits of $p$ and $C_p$ ($\alpha_p$ and $\alpha_{C_p}$

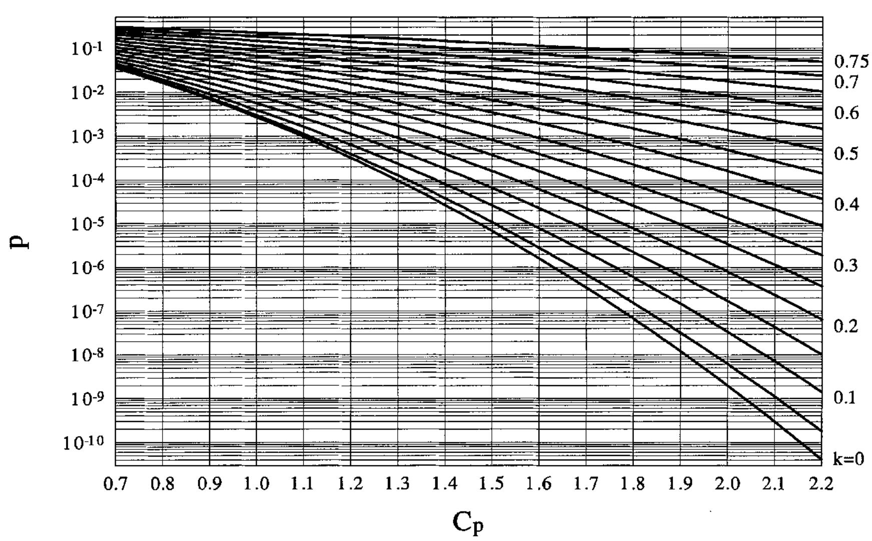


Figure 2. Behaviour of $C_p$, $k$ and $p$ on log scale

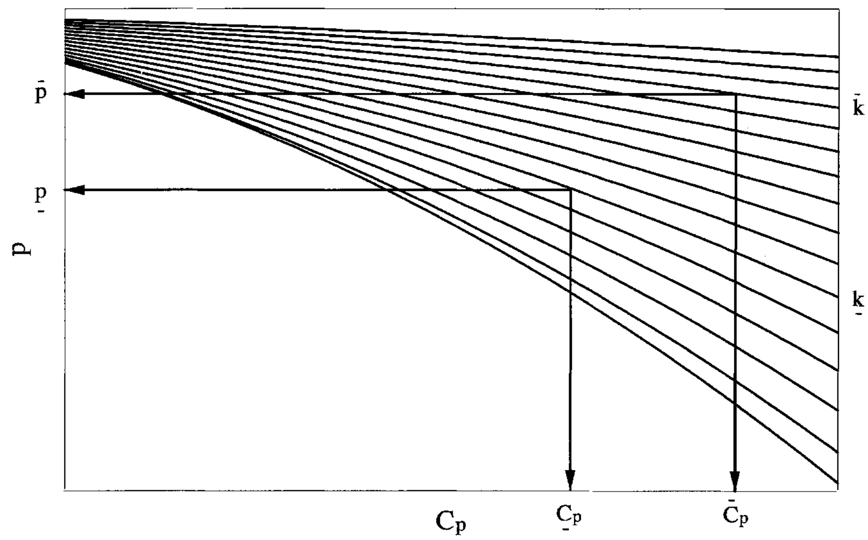


Figure 3. Graphical representation of (8) and (9)

respectively), we can then determine the minimum confidence level of the resulting confidence limits for $k$ by invoking the Bonferroni inequality. As commented by Kotz and Johnson,[10] such Bonferroni inequalities are usually overly conservative.

This conservatism originates from the large sampling variability associated with $p$. For a stable, well-established process we could assume that $p$ is known. This leads us to propose the following modifications to (8) and (9):

$$\underline{k} = \{k : L\ (k : p, \underline{C}_p) = 0\} \tag{10}$$

$$\bar{k} = \{k : L\ (k : p, \bar{C}_p) = 0\} \tag{11}$$

This assumption is not unrealistic, because the yields for most established industrial processes have been monitored closely. Graphically, the confidence limits for $k$ can be obtained by moving along the contour of constant $p$ as $C_p$ varies from $\underline{C}_p$ to $\bar{C}_p$. This is shown graphically in Figure 4.

Projecting horizontally across at the points where $\underline{C}_p$ to $\bar{C}_p$ intersect the contour of $\underline{p}$, we will arrive at the estimated values of $k$ and $\bar{k}$. From Figure 4 the estimated $\underline{k}$ then corresponds to $\underline{C}_p$ and $\bar{k}$ corresponds to $\bar{C}_p$. As shown in Figure 5, a rectangular box can also be drawn with ($\bar{C}_p, p, \bar{k}$) and ($\underline{C}_p, p, \underline{k}$) at the two opposite extreme corners. Hence the solution to (10) and (11) would simply imply locating the two corners of the postulated box.

## CONFIDENCE LIMITS FOR $C_{pk}$

The index

$$C_{pk} = \frac{\min(USL - \mu, \mu - LSL)}{3\sigma} \tag{12}$$

relates the scaled distance between the process mean ($\mu$) and the closest specification limit. $C_{pk}$ can also be written in terms of $k$ as

$$C_{pk} = (1 - k)C_p \tag{13}$$

Having established the confidence limits for $k$, we propose three approximate two-sided confidence limits for $C_{pk}$ by assuming that

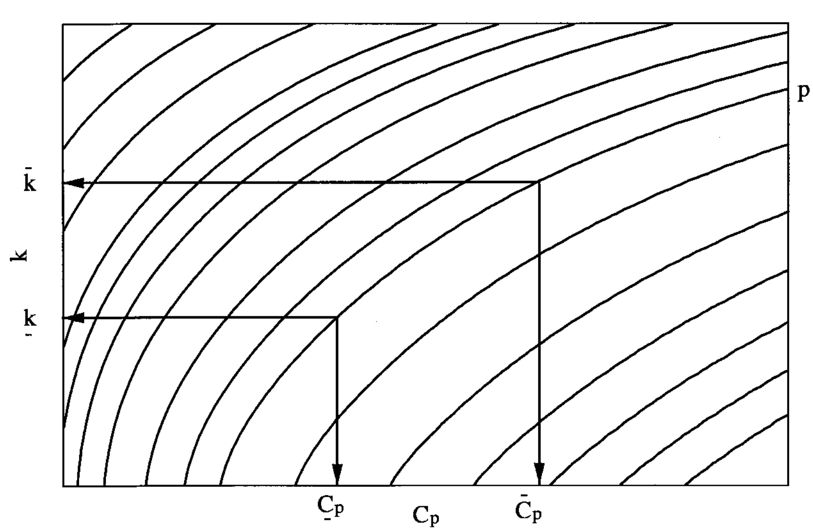


Figure 4. Graphical representation of (10) and (11)

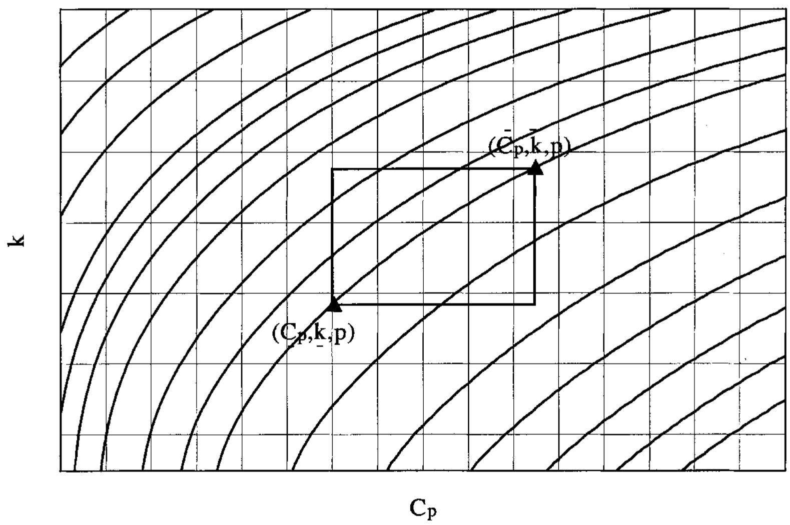


Figure 5. Tolerance box in computing upper and lower confidence limits of $k$

(a) variability in $C_p$ dominates while that of $k$ is not significant
(b) variability in $k$ dominates while that of $C_p$ is not significant
(c) variability in both $k$ and $C_p$ is significant.

*Case (a). Variability in $C_p$ dominates while that of k is not significant*

Since $C_p$ pertains to the process spread and $k$ to the shift in process mean, this case represents the practical scenario of a long-run and mature production process. In this context one has a wealth of historical data available to predict or estimate the shift in process mean accurately. With well-designed statistical experiments and process control it is possible to control the process mean at the desired target. It is the inherent variability in the process spread that is hard to eliminate; thus it dominates in the total variation of $C_{pk}$. Therefore it is appropriate to assume Case (a) when one samples from a long-run and mature production with process mean well under control. From (13) the corresponding confidence limits for $C_{pk}$ are given by

$$[\underline{C}_{pk1}, \overline{C}_{pk1}] = [(1-k)\underline{C}_p,\ (1-k)\overline{C}_p] \tag{14}$$

For a 95% confidence limit for $C_{pk1}$ we set the significance level of $C_p$ at $\alpha_{1,C_p} = 0{\cdot}05$.

*Case (b). Variability in k dominates while that of $C_p$ is not significant*

This case reflects the situation where one samples from a short-run production process during start-up or pilot run. The limited amount of data available for sampling permits only a snapshot of the process at the instant when sampling is conducted. Since the typical adjustability of the process mean is higher than that of the inherent process spread, this instant snapshot is more likely to give a better and more accurate picture of the process spread than of the process mean in the short run. Variability in the process mean or index $k$ will then be of major concern in this case when assessing the sampling variability in the index $C_{pk}$. The second type of confidence limit for $C_{pk}$ is given by

$$[\underline{C}_{pk2}, \overline{C}_{pk2}] = [(1-\overline{k})C_p,\ (1-\underline{k})C_p] \tag{15}$$

For a 95% confidence limit for $C_{pk2}$ we set the significance level of $k$ at $\alpha_{2,k} = 0{\cdot}05$.

*Case (c). Variability in both k and $C_p$ is significant*

In this case we combine the confidence limit statements of $k$ and $C_p$ using the Bonferroni inequality to give conservative confidence limits for the actual joint confidence level of $C_{pk}$. This is denoted by $[\underline{C}_{pk3}, \overline{C}_{pk3}]$, where

$$[\underline{C}_{pk3}, \overline{C}_{pk3}] = [(1-\overline{k})\underline{C}_p,\ (1-\underline{k})\overline{C}_p] \tag{16}$$

The significance levels of $C_p$ and $k$ are set at $\alpha_{3,C_p} = 0{\cdot}025$ and $\alpha_{3,k} = 0{\cdot}025$ respectively, so that through the Bonferroni inequality the joint confidence level of $C_{pk3}$ will be at least 95%, since $(1-\alpha_{C_{pk3}}) \geq 1-\alpha_{3,C_p}-\alpha_{3,k} = 0{\cdot}95$. Though this may represent the actual sampling scenario in most production processes where variability in $C_p$ and $k$ co-exists to contribute to the total variation in $C_{pk}$, the Bonferroni limits are conservative.

## A SIMULATION STUDY

To compare the performance of the proposed approximate confidence limits, a series of simulations was undertaken. Simulations each consisting of 10,000 random samples of size $n = 50$ were conducted to investigate the coverage probability and average width of the confidence limits. The values

Table I. Coverage probability and average confidence limit width from simulation at 95% confidence level ($C_p = 1{\cdot}0$, $n = 50$)

| | $k = 0{\cdot}01$ | $k = 0{\cdot}03$ | $k = 0{\cdot}05$ | $k = 0{\cdot}07$ | $k = 0{\cdot}1$ | $k = 0{\cdot}2$ | $k = 0{\cdot}3$ | $k = 0{\cdot}5$ | $k = 0{\cdot}7$ |
|---|---|---|---|---|---|---|---|---|---|
| $\underline{C}_{pk1}$ | 0·79446 | 0·77841 | 0·76236 | 0·74361 | 0·72223 | 0·64199 | 0·56174 | 0·40124 | 0·24075 |
| $\overline{C}_{pk1}$ | 1·18546 | 1·16121 | 1·13727 | 1·11333 | 1·07741 | 0·95770 | 0·83799 | 0·59856 | 0·35914 |
| Width | 0·39100 | 0·38280 | 0·37941 | 0·36702 | 0·35518 | 0·31571 | 0·27625 | 0·19732 | 0·11839 |
| Cov. prob. | 0·9422 | 0·9329 | 0·9279 | 0·9166 | 0·9077 | 0·9075 | 0·8877 | 0·8408 | 0·7009 |
| $\underline{C}_{pk2}$ | 0·77446 | 0·77129 | 0·76518 | 0·7565 | 0·7397 | 0·66632 | 0·58443 | 0·41766 | 0·25060 |
| $\overline{C}_{pk2}$ | 1·00000 | 1·00000 | 1·00000 | 0·99999 | 0·9999 | 0·99999 | 1·0000 | 0·62455 | 0·37391 |
| Width | 0·22554 | 0·22871 | 0·23482 | 0·24349 | 0·26029 | 0·33367 | 0·41557 | 0·20689 | 0·12331 |
| Cov. prob. | 0·6123 | 0·6220 | 0·6646 | 0·6997 | 0·7642 | 0·9147 | 0·9400 | 0·8483 | 0·7053 |
| $\underline{C}_{pk3}$ | 0·58637 | 0·58397 | 0·57935 | 0·57279 | 0·56009 | 0·50455 | 0·44255 | 0·31626 | 0·18976 |
| $\overline{C}_{pk3}$ | 1·22735 | 1·22735 | 1·22735 | 1·22735 | 1·22735 | 1·22735 | 1·22735 | 0·79377 | 0·47445 |
| Width | 0·64098 | 0·64338 | 0·64800 | 0·65456 | 0·66726 | 0·72280 | 0·7848 | 0·47751 | 0·28469 |
| Cov. prob. | 0·9816 | 0·9838 | 0·9857 | 0·9901 | 0·9944 | 0·9999 | 1·0000 | 0·9995 | 0·9800 |

$USL = 40{\cdot}0$, $LSL = 10{\cdot}0$ and $m = 25{\cdot}0$ were used in all simulations. In each run, $\bar{x}$ and $s$ were obtained from the 50 values generated from a normal process; $k$, $C_p$ and $C_{pk}$ were calculated from $\bar{x}$ and $s$. A 95% confidence limit of $C_{pk}$ was constructed using each of the three methods discussed in Cases (a)–(c). Each run was then replicated 10,000 times. The coverage probabilities could then be compared with the expected value determined by the significance level for $C_p$ and $k$.

The results obtained are given in Tables I–IV for $C_p = 1{\cdot}0$, 1·5, 2·0 and 2·5 respectively. The value of $\mu$ is varied in each simulation run (column), so that $k$ is set at 0·01, 0·03, 0·05, 0·07, 0·1, 0·2, 0·3, 0·5 and 0·7. The significance level for $C_p$ and $k$ has been set in such a way that the expected coverage probability for $C_{pk1}$, $C_{pk2}$ and $C_{pk3}$ is 0·95. All simulations were run on a DEC 4000 with random number generation and non-linear optimization accomplished using IMSL subroutines.

As expected, the coverage probability for Case (c) is always greater than the expected value of 0·95 for all values of $C_p$, $n$ and $k$. Its average confidence limit width is always greater than those of Cases (a) and (b). This conservative property is typical of confidence limits constructed using the Bonferroni inequality. However, the greatest interest in this study is to examine the performance of $C_{pk1}$, $C_{pk2}$ and $C_{pk3}$ for different values of $k$ and to determine the appropriate type of confidence limit to be used over different ranges of $k$.

In practice the confidence limits given by $C_{pk3}$ may be too conservative. An alternative is to propose less conservative confidence limits which are easy to construct and yet have coverage probability close to the desired value. Though not included in this paper, we have tables giving simulation results consisting of $k$- and $C_p$-values at finer resolution with sample size $n = 100$. These results are comparable and consistent with those reported here. It is important to note from Tables I–IV that the coverage probability of $C_{pk1}$ is close to the desired level for $k$ ranging from 0·01 to 0·1. This coverage probability approaches 0·95 as the value of $C_p$ increases. The average width of the $C_{pk1}$ confidence interval decreases as $k$ increases. However, this width increases as $C_p$ increases.

For $k$ ranging from 0·2 to 0·5, $C_{pk2}$ outperforms $C_{pk1}$ and $C_{pk3}$ in that the coverage probability values are closest to the targeted level. Similar to the trend exhibited by $C_{pk1}$, the coverage probability of $C_{pk2}$ approaches 0·95 as $C_p$ increases. The average confidence limit width decreases as $k$ increases but increases as $C_p$ increases.

Table II. Coverage probability and average confidence limit width from simulation at 95% confidence level ($C_p = 1{\cdot}5$, $n = 50$)

| | $k = 0{\cdot}01$ | $k = 0{\cdot}03$ | $k = 0{\cdot}05$ | $k = 0{\cdot}07$ | $k = 0{\cdot}1$ | $k = 0{\cdot}2$ | $k = 0{\cdot}3$ | $k = 0{\cdot}5$ | $k = 0{\cdot}7$ |
|---|---|---|---|---|---|---|---|---|---|
| $\underline{C}_{pk1}$ | 1·19169 | 1·16761 | 1·14354 | 1·11946 | 1·08335 | 0·96298 | 0·84261 | 0·60186 | 0·36112 |
| $\overline{C}_{pk1}$ | 1·77773 | 1·74182 | 1·70590 | 1·66999 | 1·61612 | 1·43655 | 1·25698 | 0·89784 | 0·53871 |
| Width | 0·58604 | 0·57421 | 0·56236 | 0·55053 | 0·53277 | 0·47357 | 0·41437 | 0·29598 | 0·17759 |
| Cov. prob. | 0·9436 | 0·9401 | 0·9331 | 0·9238 | 0·9243 | 0·9232 | 0·9161 | 0·8926 | 0·8179 |
| $\underline{C}_{pk2}$ | 1·21004 | 1·20026 | 1·18328 | 1·1621 | 1·12676 | 1·00239 | 0·87710 | 0·62560 | 0·37590 |
| $\overline{C}_{pk2}$ | 1·50000 | 1·50000 | 1·50000 | 1·50000 | 1·50000 | 1·50000 | 1·31208 | 0·92460 | 0·56076 |
| Width | 0·28996 | 0·29974 | 0·31672 | 0·33790 | 0·37324 | 0·49761 | 0·43498 | 0·29810 | 0·18486 |
| Cov. prob. | 0·5563 | 0·5900 | 0·6433 | 0·6948 | 0·7816 | 0·9368 | 0·9292 | 0·8971 | 0·8231 |
| $\underline{C}_{pk3}$ | 0·91627 | 0·90887 | 0·89601 | 0·87998 | 0·85322 | 0·75903 | 0·66416 | 0·4744 | 0·28464 |
| $\overline{C}_{pk3}$ | 1·84103 | 1·84103 | 1·84103 | 1·84103 | 1·84102 | 1·84103 | 1·67355 | 1·84102 | 0·71142 |
| Width | 0·92476 | 0·93216 | 0·94502 | 0·96105 | 0·98780 | 1·08200 | 1·00939 | 1·36662 | 0·42678 |
| Cov. prob. | 0·9772 | 0·9834 | 0·9842 | 0·9920 | 0·9967 | 0·9998 | 0·9997 | 1·0000 | 0·9973 |

Table III. Coverage probability and average confidence limit width from simulation at 95% confidence level ($C_p = 2{\cdot}0$, $n = 50$)

| | $k = 0{\cdot}01$ | $k = 0{\cdot}03$ | $k = 0{\cdot}05$ | $k = 0{\cdot}07$ | $k = 0{\cdot}1$ | $k = 0{\cdot}2$ | $k = 0{\cdot}3$ | $k = 0{\cdot}5$ | $k = 0{\cdot}7$ |
|---|---|---|---|---|---|---|---|---|---|
| $\underline{C}_{pk1}$ | 1·58892 | 1·55682 | 1·52472 | 1·49262 | 1·44447 | 1·28397 | 1·12348 | 0·80248 | 0·48149 |
| $\overline{C}_{pk1}$ | 2·37031 | 2·32242 | 2·27454 | 2·22665 | 2·15483 | 1·91540 | 1·67598 | 1·19713 | 0·71828 |
| Width | 0·78139 | 0·76560 | 0·74982 | 0·73403 | 0·71036 | 0·63143 | 0·55250 | 0·39465 | 0·23679 |
| Cov. prob. | 0·9445 | 0·9330 | 0·9317 | 0·9316 | 0·9331 | 0·9291 | 0·9279 | 0·9175 | 0·8760 |
| $\underline{C}_{pk2}$ | 1·63605 | 1·61573 | 1·58597 | 1·55345 | 1·50357 | 1·33653 | 1·16947 | 0·83533 | 0·50120 |
| $\overline{C}_{pk2}$ | 2·0000 | 2·0000 | 2·0000 | 2·0000 | 2·0000 | 2·00000 | 1·74479 | 1·24613 | 0·74768 |
| Width | 0·36395 | 0·38427 | 0·41403 | 0·44655 | 0·49643 | 0·66347 | 0·57532 | 0·41080 | 0·24648 |
| Cov. prob. | 0·5198 | 0·5784 | 0·6338 | 0·7028 | 0·7899 | 0·9413 | 0·9381 | 0·9227 | 0·8839 |
| $\underline{C}_{pk3}$ | 1·23886 | 1·22347 | 1·20094 | 1·17631 | 1·13854 | 1·01206 | 0·88555 | 0·63254 | 0·37952 |
| $\overline{C}_{pk3}$ | 2·45470 | 2·45470 | 2·45470 | 2·45470 | 2·45470 | 2·4547 | 2·21473 | 1·58093 | 0·94856 |
| Width | 1·21584 | 1·23123 | 1·25376 | 1·27839 | 1·31616 | 1·44264 | 1·32918 | 0·94839 | 0·56904 |
| Cov. prob. | 0·9772 | 0·9796 | 0·9879 | 0·9929 | 0·9954 | 0·9999 | 0·9998 | 0·9996 | 0·9997 |

Table IV. Coverage probability and average confidence limit width from simulation at 95% confidence level ($C_p = 2{\cdot}5$, $n = 50$)

| | $k = 0{\cdot}01$ | $k = 0{\cdot}03$ | $k = 0{\cdot}05$ | $k = 0{\cdot}07$ | $k = 0{\cdot}1$ | $k = 0{\cdot}2$ | $k = 0{\cdot}3$ | $k = 0{\cdot}5$ | $k = 0{\cdot}7$ |
|---|---|---|---|---|---|---|---|---|---|
| $\underline{C}_{pk1}$ | 1·98614 | 1·94602 | 1·90590 | 1·86577 | 1·80559 | 1·60496 | 1·40434 | 1·0031 | 0·60186 |
| $\overline{C}_{pk1}$ | 2·96289 | 2·90303 | 2·84317 | 2·78332 | 2·69353 | 2·39425 | 2·09497 | 1·49641 | 0·89784 |
| Width | 0·97675 | 0·95701 | 0·93727 | 0·91755 | 0·88794 | 0·78929 | 0·69063 | 0·49331 | 0·29598 |
| Cov. prob. | 0·9450 | 0·9365 | 0·9349 | 0·9342 | 0·9323 | 0·9314 | 0·9278 | 0·9214 | 0·8970 |
| $\underline{C}_{pk2}$ | 2·05724 | 2·02449 | 1·98379 | 1·94214 | 1·87950 | 1·67067 | 1·46183 | 1·04417 | 0·62650 |
| $\overline{C}_{pk2}$ | 2·50000 | 2·50000 | 2·50000 | 2·50000 | 2·50000 | 2·49999 | 2·18074 | 1·55767 | 0·93460 |
| Width | 0·44276 | 0·47551 | 0·51621 | 0·55786 | 0·62050 | 0·82932 | 0·71891 | 0·51350 | 0·30810 |
| Cov. prob. | 0·5157 | 0·5652 | 0·6390 | 0·7030 | 0·7925 | 0·9428 | 0·9393 | 0·9316 | 0·9052 |
| $\underline{C}_{pk3}$ | 1·5578 | 1·53300 | 1·50218 | 1·47064 | 1·42321 | 1·26507 | 1·10694 | 0·79067 | 0·47440 |
| $\overline{C}_{pk3}$ | 3·06837 | 3·06837 | 3·06837 | 3·06837 | 3·06837 | 3·06837 | 2·76674 | 1·97617 | 1·18570 |
| Width | 1·51057 | 1·53537 | 1·56619 | 1·59773 | 1·64516 | 1·8033 | 1·65980 | 1·1855 | 0·71130 |
| Cov. prob. | 0·9728 | 0·9797 | 0·9882 | 0·9920 | 0·9958 | 0·9996 | 1·00000 | 0·9996 | 0·9997 |

By examining the performance of the three different types of confidence limits for different $k$-values, we could determine the ‘breakpoint’ where the contribution of $k$ to the total variation becomes dominant in $C_p$ and can be assumed to be the sole $C_{pk}$ total variation contributor. This occurs for $k$ ranging from 0·2 to 0·5. The ‘breakpoint’ is consistently located at $k = 0{\cdot}2$ even when $C_p$ changes. A similar observation is reported for simulation at $n = 100$. Hence the ‘breakpoint’ is robust to the values of $C_p$ and $n$.

Table V serves to determine the appropriate type of confidence limit for different ranges of $k$. For $k > 0{\cdot}5$ we should not concern ourselves with obtaining the confidence limits for $C_{pk}$, since it is highly questionable whether the process itself is in statistical control. Such a large shift in the mean of the process is likely to have been detected by the control chart.

## ILLUSTRATIVE EXAMPLES

We use the following examples to demonstrate the usage of AM confidence limits for $C_{pk}$ under two different scenarios.

*Example* 1

A company has a well-established process, say process X1, to supply a certain product for its customer. To ensure low incoming defects, the customer specifies a minimum $C_{pk}$-value of 1·8, with a lower confidence limit of 1·5 to cater for sampling variability. A process capability study conducted on process X1 reveals the following information: $n = 100$, $USL = 30$, $LSL = 12$, $\bar{x} = 21{\cdot}27$ and $s = 1{\cdot}5$. Assume that the desired confidence level for all related calculations is 95% ($\alpha = 0{\cdot}05$).

Substituting these data in (1)–(4) will give

Table V. Practitioner’s guide

| | $k < 0{\cdot}1$ | $k = 0{\cdot}1$–$0{\cdot}2$ | $k = 0{\cdot}2$–$0{\cdot}5$ | $k > 0{\cdot}5$ |
|---|---|---|---|---|
| $\underline{C}_{pk}$ | $\underline{C}_{pk1}$ | $\underline{C}_{pk1}$ or $\underline{C}_{pk2}$ | $\underline{C}_{pk2}$ | Adjust process average? |
| $\overline{C}_{pk}$ | $\overline{C}_{pk1}$ | $\overline{C}_{pk1}$ or $\overline{C}_{pk2}$ | $\overline{C}_{pk2}$ | Adjust process average? |

$$\hat{C}_p = \frac{30 - 12}{6 \times 1{\cdot}5} = 2{\cdot}0$$

$$m = \tfrac{1}{2}(30 + 12) = 21$$

$$\hat{k} = \frac{|21 - 21{\cdot}27|}{(30 - 12)/2} = 0{\cdot}03$$

$$\hat{C}_{pk} = (1 - 0{\cdot}03)(2{\cdot}0) = 1{\cdot}94$$

Since $\hat{k}$ is less than 0·1, Table V recommends $\underline{C}_{p_{k1}}$ as the lower AM confidence limit for $C_{pk}$. The $k$-value is reasonable in this case for a well-established and optimized process.

From (6), $\underline{C}_p$ is given as

$$\underline{C}_p = \frac{\chi^2_{100-1,0{\cdot}05/2}}{\sqrt{100 - 1}} \times 2{\cdot}0 = 1{\cdot}684$$

From (14), $\underline{C}_{pk1}$ is calculated as

$$\underline{C}_{pk1} = (1 - 0{\cdot}03)(1{\cdot}684) = 1{\cdot}633$$

We can thus safely claim to the customer that process X1 meets the minimum $C_{pk}$-value and lower confidence limit value.

*Example* 2

Another process X2 is a new process acquired by a supplier. We wish to qualify the process by running a test production. The customer specifies a minimum $C_{pk}$-value of 1·3, with a lower confidence limit of 1·0. Since X2 is a new process, it is yet to be optimized and controlled for production. A process capability study conducted for X2 records the following data: $n = 50$, $USL = 20{\cdot}8$, $LSL = 10$, $\bar{x} = 17{\cdot}2$ and $s = 1{\cdot}2$.

Substituting these data in (1)–(4) will give

$$\hat{C}_p = \frac{20{\cdot}8 - 10}{6 \times 1{\cdot}2} = 1{\cdot}5$$

$$m = \tfrac{1}{2}(20{\cdot}8 + 10) = 15{\cdot}4$$

$$\hat{k} = \frac{|15{\cdot}4 - 17{\cdot}2|}{(20{\cdot}8 - 10)/2} = 0{\cdot}3$$

$$\hat{C}_{pk} = (1 - 0{\cdot}3)(1{\cdot}5) = 1{\cdot}05$$

Since $\hat{k}$ falls between 0·2 and 0·5, Table V recommends $\underline{C}_{pk2}$ as the lower AM confidence limit for $C_{pk}$. The expression for $\underline{C}_{pk2}$ in (15) requires the determination of $\bar{k}$.

From Figure 4 we can determine $\bar{k}$ by knowing $p$ and $\bar{C}_p$, which are easily calculated from (5) and (6) respectively, i.e.

$$p = \Phi[-3(1 + 0{\cdot}3)1{\cdot}5] + \Phi[-3(1 - 0{\cdot}3)1{\cdot}5]$$
$$= 8{\cdot}164 \times 10^{-4}$$

$$\bar{C}_p = \frac{\chi^2_{50-1,1-0{\cdot}05/2}}{\sqrt{50 - 1}} \times 1{\cdot}5 = 1{\cdot}796$$

Then $\bar{k}$ can be obtained either by estimating it directly from Figure 4 as 0·40 or by solving (11) as below using some numerical solvers:

$$\bar{k} = \{k : L(k : p = 8{\cdot}164 \times 10^{-4}, \bar{C}_p = 1{\cdot}796) = 0\}$$
$$= 0{\cdot}415$$

From (15), using the exact $\bar{k}$-value, $\underline{C}_{pk2}$ is determined to be

$$\underline{C}_{pk2} = (1 - 0{\cdot}415)(1{\cdot}5) = 0{\cdot}878$$

Clearly, we have little faith in process X2 to meet the customer's quality requirement. We can also proceed to calculate the confidence limits for the fraction non-conforming for this particular example. It is probable that the customer will be most interested in the value $\bar{p}$.

From Figure 2 we can locate $\bar{p}$ by knowing $\bar{k}$ and $C_p$ (recall that $\underline{C}_{pk2} = (1 - \bar{k})C_p$), which were determined earlier as 0·415 and 1·5 respectively.

Then $\bar{p}$ can be determined either by estimating it directly from Figure 2 as approximately $4{\cdot}0 \times 10^{-3}$ or by calculating it from (5) using $\bar{k}$ and $C_p$ as

$$\bar{p} = \Phi[-3(1 + 0{\cdot}415)1{\cdot}5] + \Phi[-3(1 - 0{\cdot}415)1{\cdot}5]$$
$$= 4{\cdot}24 \times 10^{-3}$$

## COMPARISON WITH BOOTSTRAP CONFIDENCE LIMITS

We have thus far arrived at some approximate confidence limits that are computationally easy to obtain and provide adequate protection in terms of their coverage probability. In this section we compare the proposed approximate method (AM) confidence limits with three non-parametric bootstrap confidence limits for $C_{pk}$, i.e. the standard bootstrap (SB), the percentile bootstrap (PB) and the biased-corrected percentile bootstrap (BCPB) reported by Franklin and Wasserman.[9]

The performance comparison involves a series of simulations. The values $USL = 61$, $LSL = 40$ and target $m = 50{\cdot}5$ were used for all simulations. The six defined parameter values used in the simulation study are given in Table VI. These values were chosen to represent processes that vary from 'vary capable' to 'not capable'. To calculate the bootstrap

Table VI. Six values of parameters used in simulation study

| $\mu$ | $\sigma$ | $k$ | $C_p$ | $C_{pk}$ |
|---|---|---|---|---|
| 50 | 2 | 0·0476 | 1·667 | 1·565 |
| 52 | 2 | 0·1429 | 1·500 | 0·971 |
| 50 | 3 | 0·0476 | 1·111 | 1·107 |
| 52 | 3 | 0·1429 | 1·000 | 0·825 |
| 50 | 3·7 | 0·0476 | 0·901 | 0·913 |
| 52 | 3·7 | 0·1429 | 0·811 | 0·735 |

confidence limits for each combination, a sample of size $n = 30$, 50 or 75 was drawn and 1000 bootstrap resamples (each of size $n$) were drawn from that single sample. A 90% bootstrap lower confidence limit for the index $C_{pk}$ was constructed by each of the SB, PB and BCBP methods. This single simulation was then replicated 1000 times. The frequency for each of the 90% bootstrap confidence intervals containing the true $C_{pk}$-value was recorded. An average length of the bootstrap confidence limits was also calculated.

To derive the corresponding proposed AM confidence limits, we used the same parameters and number of trials (1000) as those in bootstrap for our simulations. The simulation was conducted in a similar fashion as described in the previous section. Since all the $k$-values in this simulation are less than 0·01, we used $[\underline{C}_{pk1}, \overline{C}_{pk1}]$ as our AM confidence limits for all six simulation runs. To achieve a target coverage probability of 0·90, $\alpha_{1,C_p}$ was set at 0·10. The results of the performance comparison are given in Table VII.

The SB method gives coverage probabilities consistently near the expected value of 0·90 all the time. In contrast, the AM, PB and BCBP limits have coverage probabilities lower than 0·90. All three limits tend to increase slowly towards 0·90 as $n$ increases. In general the AM gives larger coverage probabilities than PB and BCBP, except for the last two simulations. SB is the only one of the three bootstrap confidence intervals that consistently gives 0·90 coverage. The superiority of AM comes in when we examine the average width of the limits. As expected from the theory, the average width decreases as $n$ increases for all limits. The AM intervals are consistently the narrowest, followed by BCBP, PB and SB. There is an approximately 10% difference between the AM average width and the narrowest average width of the three bootstrap intervals.

The worst of the four methods was PB, which has the lowest coverage probability most of the time. It has also the largest average width, which in this respect is comparable with the SB method. The AM has the smallest average width of the four methods. The difference is largest for small sample sizes and tends to reduce as $n$ increases to 75. The AM also has slightly lower coverage probability than the SB method, except for low $C_{pk}$ ($<1·0$) where the discrepancy increases. For high $C_{pk}$-values the AM attains coverages comparable with the SB method and yet has the narrowest average widths.

Table VII. $C_{pk}$ confidence limits—performance comparison of SB, PB, BCBP and AM

| | $n = 30$ | | $n = 50$ | | $n = 75$ | |
|---|---|---|---|---|---|---|
| | Cov. prob. | Ave. width | Cov. prob. | Ave. width | Cov. prob. | Ave. width |
| ($\mu = 50$, $\sigma = 2$, $k = 0·0476$, $C_p = 1·750$, $C_{pk} = 1·667$, $C_{pk} = \underline{C}_{pk1}$) | | | | | | |
| SB | 0·894 | 0·799 | 0·895 | 0·584 | 0·899 | 0·476 |
| PB | 0·841 | 0·785 | 0·856 | 0·579 | 0·875 | 0·474 |
| BCBP | 0·852 | 0·727 | 0·864 | 0·552 | 0·878 | 0·458 |
| *AM* | *0·885* | *0·717* | *0·879* | *0·552* | *0·882* | *0·450* |
| ($\mu = 52$, $\sigma = 2$, $k = 0·1429$, $C_p = 1·750$, $C_{pk} = 1·500$, $C_{pk} = \underline{C}_{pk1}$) | | | | | | |
| SB | 0·902 | 0·729 | 0·897 | 0·545 | 0·923 | 0·434 |
| PB | 0·842 | 0·715 | 0·858 | 0·540 | 0·894 | 0·432 |
| BCBP | 0·854 | 0·651 | 0·873 | 0·510 | 0·906 | 0·416 |
| *AM* | *0·877* | *0·645* | *0·890* | *0·497* | 0·899 | *0·405* |
| ($\mu = 50$, $\sigma = 3$, $k = 0·0476$, $C_p = 1·167$, $C_{pk} = 1·111$, $C_{pk} = \underline{C}_{pk1}$) | | | | | | |
| SB | 0·901 | 0·527 | 0·890 | 0·399 | 0·893 | 0·322 |
| PB | 0·883 | 0·519 | 0·880 | 0·396 | 0·890 | 0·321 |
| BCBP | 0·880 | 0·495 | 0·880 | 0·386 | 0·877 | 0·315 |
| *AM* | *0·889* | *0·478* | *0·887* | *0·368* | *0·886* | *0·300* |
| ($\mu = 52$, $\sigma = 3$, $k = 0·1429$, $C_p = 1·167$, $C_{pk} = 1·000$, $C_{pk} = \underline{C}_{pk1}$) | | | | | | |
| SB | 0·900 | 0·506 | 0·904 | 0·378 | 0·909 | 0·306 |
| PB | 0·825 | 0·496 | 0·866 | 0·375 | 0·884 | 0·306 |
| BCBP | 0·841 | 0·457 | 0·875 | 0·356 | 0·888 | 0·295 |
| *AM* | *0·851* | *0·450* | *0·854* | *0·331* | *0·867* | *0·270* |
| ($\mu = 50$, $\sigma = 3·7$, $k = 0·0476$, $C_p = 0·946$, $C_{pk} = 0·901$, $C_{pk} = \underline{C}_{pk1}$) | | | | | | |
| SB | 0·892 | 0·432 | 0·896 | 0·329 | 0·879 | 0·269 |
| PB | 0·875 | 0·425 | 0·887 | 0·327 | 0·880 | 0·268 |
| BCBP | 0·867 | 0·413 | 0·876 | 0·322 | 0·868 | 0·265 |
| *AM* | *0·869* | *0·387* | *0·878* | *0·299* | *0·874* | *0·243* |
| ($\mu = 52$, $\sigma = 3·7$, $k = 0·1429$, $C_p = 0·946$, $C_{pk} = 0·811$, $C_{pk} = \underline{C}_{pk1}$) | | | | | | |
| SB | 0·896 | 0·429 | 0·899 | 0·317 | 0·896 | 0·256 |
| PB | 0·849 | 0·423 | 0·862 | 0·315 | 0·867 | 0·256 |
| BCBP | 0·860 | 0·395 | 0·874 | 0·301 | 0·879 | 0·248 |
| *AM* | *0·845* | *0·349* | *0·863* | *0·269* | *0·865* | *0·219* |

## CONCLUSIONS

We have proposed an approximate method to derive the confidence limits for $C_{pk}$. The method is easier to compute compared with other methods such as bootstrap and non-central $t$-integration. A graphical method was used to show the relationship between $C_p$, $p$ and $k$. Based on the graphs, we derived the confidence limits of $k$ and used the tolerance box to relate the confidence limits of $C_p$, $p$ and $k$. We then evaluated three approximate $C_{pk}$ confidence limits, i.e. $C_{pk1}$, $C_{pk2}$ and $C_{pk3}$, using simulations and proposed a guide for practitioners to arrive at the 'approximate method' (AM) confidence limits. We also demonstrated their usage via illustrative examples. These AM limits were shown to have good coverage probabilities and short average confidence interval widths as compared with three bootstrap confidence limits. Their performance is robust to the sample size $n$ and capability index $C_p$.

## REFERENCES

1. V. E. Kane, 'Process capability indices', *J. Qual. Technol.*, **18**, 41–52 (1986).
2. L. P. Sullivan, 'Reducing variability: a new approach to quality', *Qual. Prog.*, July, pp. 15–21 (1984).
3. D. C. Montgomery, *Introduction to Statistical Quality Control*, Wiley, New York, 1985.
4. C.W. Wang, 'Letters to the Editor', *Qual. Reliab. Engng. Int.*, **6**, 312–316 (1990).
5. R. Kushler and P. Hurley, 'Confidence bounds for capability indices', *J. Qual. Technol.*, **24**, 188–195 (1992).
6. Y. M. Chou, D. B. Owen and A. S. A. Borrego, 'Lower confidence limits on process capability indices', *J. Qual. Technol.*, **22**, 223–229 (1990).
7. N.F. Zhang, G.A. Stenback and D.M. Wardrop, 'Interval estimation of process capability index $C_{pk}$' *Commun. Stat.—Theory Meth.*, **19**, 4455–4470 (1990).
8. A. F. Bissell, 'How reliable is your capability index?', *Appl. Stat.*, **39**, 331–340 (1990).
9. L. A. Franklin and G. Wasserman, 'Bootstrap confidence interval estimates of $C_{pk}$: an introduction', *Commun. Stat.—Simul. Comput.*, **20**, 231–242 (1991).
10. S. Kotz and N. L. Johnson, *Process Capability Indices*, Chapman and Hall, London, 1993.

*Authors' biographies:*

**Loon Ching Tang** is a Senior Lecturer in the Department of Industrial and Systems Engineering, National University of Singapore. He received his M.Sc. and Ph.D. in Operations Research from Cornell University and his B.Eng. (1st class) and M.Eng. in Mechanical Engineering and Industrial Engineering respectively from the National University of Singapore. His current research and consulting focus on applications of probability and statistics to industrial problems, particularly in the areas of quality, reliability, simulation and operations planning.

**Sue Ee Than** is a research scholar in the Department of Industrial and Systems Engineering, National University of Singapore. He received his B.Eng. (1st class) in Mechanical Engineering from the National University of Singapore. An ex-ASEAN and Malaysia Tunku Abdul Rahman scholar, his main research interests are in statistical quality engineering and simulation.

**Beng Wah Ang** is an Associate Professor in the Department of Industrial and Systems Engineering, National University of Singapore. His current research interests include systems modelling and forecasting, and quality engineering and management. He teaches and consults in the areas of statistical quality control and engineering statistics. He received his Ph.D. from the University of Cambridge and his B.Sc. in Applied Physics from Nanyang University, Singapore.