

Exact expression for the expectation of estimated C_{pk} based on control chart information and the corresponding process capability control charts

Moutushi Chatterjee

To cite this article: Moutushi Chatterjee (2017) Exact expression for the expectation of estimated C_{pk} based on control chart information and the corresponding process capability control charts, Journal of Statistical Computation and Simulation, 87:5, 1025-1047, DOI: [10.1080/00949655.2016.1243684](https://doi.org/10.1080/00949655.2016.1243684)

To link to this article: <http://dx.doi.org/10.1080/00949655.2016.1243684>



Published online: 16 Oct 2016.



Submit your article to this journal [↗](#)



Article views: 9



View related articles [↗](#)



View Crossmark data [↗](#)

Exact expression for the expectation of estimated C_{pk} based on control chart information and the corresponding process capability control charts

Moutushi Chatterjee

^aSQC & OR Unit, Indian Statistical Institute, Kolkata, India; ^bDepartment of Statistics, Lady Brabourne College, Kolkata, India

ABSTRACT

Among all the process capability indices, C_{pk} is the most popular among the practitioners, when the underlying quality characteristic follow normal distribution. Due to its complicated expression, properties of its plug-in estimator have been of particular interest to several researchers both from classical and Bayesian statistical arena. In the present article, we have proposed plug-in estimators of C_{pk} based on the $\bar{X} - R$ and $\bar{X} - S$ chart information under the assumption that each of the samples are drawn independently. We have also derived the expressions for expected values of these estimators and the associated bias. These estimators are found to work similar to the Bayesian-like estimators without depending upon any historical information. The corresponding process capability control charts have also been designed to enable continuous assessment of a process. Finally, a simulated example and two real-life examples have been discussed to supplement the theory developed in this article.

ARTICLE HISTORY

Received 1 December 2015
Accepted 27 September 2016

KEYWORDS

Bias; Bayesian-like estimator; continual process assessment; control chart; C_{pk} ; plug-in estimator; process capability control chart

2010 MATHEMATICS

SUBJECT CLASSIFICATION
62E15; 62F10; 62F15; 62F25; 62Q05

1. Introduction

Process capability index (PCI) is one of the most versatile statistical tool to assess the capability of a process to produce items within the pre-assigned specification limits.

Under the assumption of normality of the distribution of the quality characteristic under consideration, the four classical PCIs for univariate processes with bilateral specification limits are

$$\begin{aligned} C_p &= \frac{USL - LSL}{6\sigma}, \\ C_{pk} &= \frac{d - |\mu - M|}{3\sigma}, \\ C_{pm} &= \frac{d}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \\ C_{pmk} &= \frac{d - |\mu - M|}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \end{aligned} \quad (1)$$

where USL and LSL are, respectively, the upper and lower specification limits of a process, $d = (USL - LSL)/2$, $M = (USL + LSL)/2$ and T is the target of the process (refer [1,2]).

Among these PCIs, C_p measures the potential capability of a process, that is, the capability level achievable by a process when it is perfectly centred on target; C_{pk} takes into account both the process centring as well as the process dispersion with respect to the concerned quality characteristic; C_{pm} measures the proximity of the process centring towards the target T and C_{pmk} is the third generation PCI (refer [1]) which is defined combining C_{pk} and C_{pm} . Among these, C_{pk} is the most popular PCI among the practitioners in manufacturing industries.

Like most of the PCIs available in the literature, C_{pk} is also defined as a function of the parameters of the quality characteristics, namely μ and σ [vide equation (1)]. Since μ and σ are often unobservable in practice, the true value of C_{pk} is also unobservable. This necessitates defining its plug-in (natural) estimator as

$$\hat{C}_{pk} = \frac{d - |\bar{X} - M|}{3S}, \quad (2)$$

where $\bar{X} = (1/n) \sum_{i=1}^n X_i$ is the sample mean; $S = \sqrt{(1/(n-1)) \sum_{i=1}^n (X_i - \bar{X})^2}$ is the sample standard deviation and n is the sample size.

Although, the common industrial practice is to draw conclusion about the so-called capability or incapability of a process based on $\hat{C}_{pk}$ value itself; since it is based on mere sample observations, often such $\hat{C}_{pk}$ values are subjected to the perils of sampling fluctuations. Hence, the statistical properties of $\hat{C}_{pk}$ need to be studied extensively. However, the existence of the modulus function in the definition of $\hat{C}_{pk}$ makes such study somewhat difficult and hence has drew substantial attention of many eminent statisticians (refer [2,3] and the references there-in). Novoa and Leon [4] derived the statistical distribution of $\hat{C}_{pk}$, based on the concept of folded normal distribution. The testing of hypothesis problems related to $\hat{C}_{pk}$ have been studied by Chen and Hsu [5] and Lin [6] among others.

However, it has been observed time and again that, inference drawn on process capability based on single sample information is often subject to high degree of sampling fluctuations. Hence a more plausible alternative is to use subsample information to define the plug-in estimator of C_{pk} . The properties of such estimators of C_{pk} have been studied by Lin and Sheen [7,8].

Due to its huge popularity among practitioners, C_{pk} has found substantial application in developing various sampling inspection plans (refer [9–14] and so on) and product acceptance determination (refer [15]). Pearn and Liao [16] discussed about the procedure of assessing the capability of a process using C_{pk} , when the measurements of the concerned quality characteristic suffer from gauge measurement error.

Although, most of the studies of the distributional and inferential properties of $\hat{C}_{pk}$ are based on the so-called classical frequentist approach, ample research work have been carried in this field from Bayesian perspective as well (refer [17–20] and so on). Pearn and Chen [17,18] defined Bayesian-like estimators of C_{pk} and studied their distributional and inferential properties. Later, Lin and Sheen [7] have studied some inferential properties of the Bayesian-like estimator of C_{pk} , defined by Pearn and Chen [17], using control chart data.

Note that, proper evaluation of such Bayesian-like estimator requires adequate knowledge of historical information, like $P_r(\mu < M)$ and $P_r(\mu > M)$, of a statistically stable process. However, often, such historical data may not be always available and/or reliable due to the following reasons:

- (1) The concerned authority may be ignorant about the storage of historical data .
- (2) Even if the historical data are available, it may not be reliable for further use, due to anomaly in measuring system or erratic practice of collecting the data .
- (3) Finally, there remains ample room for favourable manipulation in the estimated C_{pk} value, by simply playing with the prior probabilities, namely $P_r(\mu < M)$ and $P_r(\mu > M)$.

In this context, while constructing the lower confidence bounds (LCB) of C_{pk} , based on classical approach and lower credible bound based on Bayesian approach, Pearn et al. [21] have suggested

using the classical approach over the Bayesian one. To be more precise, the authors observed that, the LCB computed using the classical method is always greater than its Bayesian counterpart.

Now, before assessing the capability of a process through a suitable PCI, one needs to check and establish the statistical stability of the said process through the use of appropriate control charts like $\bar{X}$ -R or $\bar{X}$ -S charts (refer [22]). Although, for a statistically stable process, the estimates of the parameters (namely μ and σ) can be obtained directly from these control charts, in practice, they are not used in the subsequent stages of capability assessment. Rather, fresh samples are drawn to estimate the same parameters – making the entire procedure highly uneconomical (refer [23]).

Pearn et al. [21] defined the plug-in estimators of C_p based on $\bar{X}$ -R and $\bar{X}$ -S control chart information and discussed about the corresponding hypothesis testing procedure. They have also suggested a step-by-step procedure to enable the practitioners to decide whether the concerned process is capable or not. Lin and Sheen [7] argued for estimating C_{pk} based on control chart information. However, they have adopted Bayesian approach requiring the knowledge of the probabilities $P_r(\mu < M)$ and $P_r(\mu > M)$, which may not always be available.

In the present article, we have defined a plug-in estimator of C_{pk} based on $\bar{X}$ -R and $\bar{X}$ -S control chart information and have derived the expressions for the corresponding expectations and the associated biases. Since this addresses the problem of volatility of the estimators based on single sample information and at the same time, does not require any historical information, it is more suitable for application in practice than the existing estimators of C_{pk} .

Interestingly, despite being the most adopted approach in practice, point estimates of PCIs may not always depict the true health of a process. Infact, being a continuous flow of activity, the performance of a process is supposed to change over the entire production cycle. This should be reflected by the associated PCIs (refer [23,24]). Unfortunately, while PCIs estimated using single sample information may not always represent the overall capability scenario of a process due to the influence of sampling fluctuation (refer [25]); PCIs estimated using multiple sample information are often prone to smooth out some important fluctuations in the capability level of a process (refer [23]).

Boyles [26] first advocated the use of process capability control charts, in the context of the PCI C_{pm} (refer Equation (1)) to capture the changes in capability level of a process over time. Later Spiring [25] redefined the control limits of this C_{pm} control chart by introducing the use of the information gathered from the corresponding $\bar{X}$ -R and $\bar{X}$ -S charts. Spiring [27] has discussed capability assessment of short-run processes using C_{pm} , based on control chart data. However, as has been observed by Chatterjee and Chakraborty [23], Spiring's [25] approach suffered from some fundamental problems in a sense that, while developing the control limits of C_{pm} control chart, Spiring [25] made use of the plug-in estimators of C_{pm} based on control chart data, while the associated distribution was that of the single sample estimator of C_{pm} .

Chatterjee and Chakraborty [23,24] fine-tuned the methodology adopted by Spiring [25] and designed process capability control charts of $C_{pu} = (USL - T)/3\sigma$, $C_{pl} = (T - LSL)/3\sigma$ and some other popular PCIs for unilateral specification limits. The authors have also enlisted some advantages of using process capability control charts over the use of the single or multiple sample-based plug-in estimators of PCIs, like (i) requirement of lesser sample size, (ii) reduction in sampling cost – making the entire procedure of process capability analysis more economical – and (iii) enabling continuous assessment of process performance.

Apart from these, Kuo [28] designed a control chart for C_p , based on R-chart information. Leung and Spiring [29] modified the control limits of C_{pm} control chart for the processes, where the assumption of normality of the underlying distribution is violated. Some other interesting properties of C_{pm} control chart have also been studied by Kuo et al. [30], Morita et al. [31] and Wu [32]. Wu [32] used the concept of target costing to establish goal control limits, based on the relationship among the loss function, PCIs and control charts. Testing process capability in the context of unilateral specification limits with subgroup information was carried out by Wu [33]. Wu [34] discussed about computing PCI values based on subsample information from Bayesian perspective.

Thus, due to its enormous practical utility, control charts have been designed for most of the popular PCIs like C_p , C_{pm} , C_{pu} , C_{pl} and so on. However, despite being one of the most widely accepted PCIs among practitioners, the case of C_{pk} has been grossly neglected in this regard, probably due to its complicated statistical distribution. Although Novoa and Leon [4] designed a control chart for C_{pk} , they did not provide any explicit form of the corresponding control limits. Moreover, similar to Spiring [25], they also did not take the contribution of the number of subgroups (say, m) into account. To address these problems, we have designed process capability control charts of C_{pk} based on the $\bar{X} - R$ and $\bar{X} - S$ chart information.

In the next section, we have defined plug-in estimators of C_{pk} using $\bar{X} - R$ and $\bar{X} - S$ chart information and have derived the corresponding expectations along with the associated biases. In Section 3, the corresponding process capability control charts of C_{pk} are designed. Some numerical examples are discussed in Section 4 to supplement the theory developed in this article. Finally, we conclude the article in Section 5 with a brief summary of the article.

2. Estimation of C_{pk} based on $\bar{X} - R$ and $\bar{X} - S$ chart information and the associated bias

Following Pearn and Kotz [2], C_{pk} can be redefined as

$$C_{pk} = \left\{ 1 - \frac{|\mu - M|}{d} \right\} \times \frac{d}{3\sigma}. \quad (3)$$

Suppose the quality characteristic under consideration is normally distributed and the concerned process is under statistical control. Also, suppose, while checking stability of the process, we have m rational subgroups (refer [22]) and from each rational subgroup, a sample of size n is drawn. Here we assume constant sample size for all the rational subgroups. So the total number of observations is $N = mn$. Also let X_{ij} is the measured value of the quality characteristic for the j th sample from the i th rational subgroup such that $X_{ij} \sim N(\mu, \sigma^2)$, for $i = 1(1)m$ and $j = 1(1)n$. Then,

- (1) $\bar{\bar{X}} = (1/N) \sum_{i=1}^m \sum_{j=1}^n X_{ij} = (1/m) \sum_{i=1}^m \bar{X}_i$ is the average of subgroup averages .
- (2) $\bar{S} = (1/m) \sum_{i=1}^m S_i$ is the average of the standard deviations over the subgroups with $S_i = \sqrt{1/(n-1) \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2}$ being the standard deviation corresponding to the i th subgroup for $i = 1(1)m$.
- (3) $\bar{R} = (1/m) \sum_{i=1}^m R_i$ is the average of the ranges over the subgroups with R_i being the range corresponding to the i th subgroup for $i = 1(1)m$.

2.1. Plug-in estimator of C_{pk} based on $\bar{X} - R$ chart information and the corresponding expectation

The plug-in (natural) estimator of C_{pk} , based on $\bar{X} - R$ chart information, can be defined, by replacing μ and σ , respectively, by $\bar{\bar{X}}$ and $\bar{R}/d_2$ in Equation (3), as

$$\hat{C}_{pk}^{(R)} = \frac{dd_2}{3} \left\{ 1 - \frac{|\bar{\bar{X}} - M|}{d} \right\} \times \frac{1}{\bar{R}}, \quad (4)$$

where d_2 is a function of the sample size n (refer [22]).

Following Lord [35] and Lin and Sheen [7], $\bar{\bar{X}}$ and $\bar{R}$ are mutually independent under the assumption of normality of the quality characteristic under consideration. Thus from Equation (4)

$$E[\hat{C}_{pk}^{(R)}] = \frac{dd_2}{3} \times \left[1 - E \left(\frac{|\bar{\bar{X}} - M|}{d} \right) \right] \times E(\bar{R}^{-1}). \quad (5)$$

We have already assumed that $X_{ij} \sim N(\mu, \sigma^2)$, which implies, $\bar{\bar{X}} \sim N(\mu, \sigma^{*2})$, where, $\sigma^{*2} = \sigma^2/N$. Hence, following Leone et al. [36], $|\bar{\bar{X}} - M|$ follows folded normal distribution with mean

$$\mu_f^* = \frac{\sigma\sqrt{2}}{\sqrt{N\pi}} \times e^{-\frac{N(\mu-M)^2}{2\sigma^2}} + (\mu - M) \times \left[1 - 2\Phi\left(-\frac{(\mu - M)\sqrt{N}}{\sigma}\right) \right], \quad (6)$$

where Φ denotes the cdf of univariate standard normal distribution and variance

$$\sigma_f^{*2} = (\mu - M)^2 + \frac{\sigma^2}{N} - \mu_f^{*2}. \quad (7)$$

Notationally, $|\bar{\bar{X}} - M| \sim FN(\mu_f^*, \sigma_f^{*2})$.

Hence,

$$E[|\bar{\bar{X}} - M|] = \mu_f^*. \quad (8)$$

Also, following Woodall and Montgomery [37], $\bar{R}/\sigma \sim (d_2^*/\sqrt{v}) \times \chi_v$ approximately, where, $v = 1/(-2 + 2\sqrt{1 + 2/m \times (d_3/d_2)^2})$ and $d_2^* = \sqrt{d_2^2 + d_3^2/m}$, d_3 being a function of n (refer [22]). Thus,

$$\begin{aligned} E(\bar{R}^{-1}) &= \left(\frac{\sqrt{v}}{\sigma d_2^*} \right) \times E(\chi_v^{-1}) \\ &= \sqrt{\frac{v}{2}} \times \frac{1}{d_2^* \sigma} \times \frac{\Gamma(\frac{v-1}{2})}{\Gamma(\frac{v}{2})}. \end{aligned} \quad (9)$$

Hence, substituting Equations (8) and (9) in Equation (5)

$$\begin{aligned} E[\hat{C}_{pk}^{(R)}] &= \frac{dd_2}{3} \times \left[1 - \frac{\mu_f^*}{d} \right] \times \sqrt{\frac{v}{2}} \times \frac{1}{d_2^* \sigma} \times \frac{\Gamma(\frac{v-1}{2})}{\Gamma(\frac{v}{2})} \\ &= \frac{dd_2}{3\sigma b_v d_2^*} \times \left(1 - \frac{\mu_f^*}{d} \right), \end{aligned} \quad (10)$$

where $b_v = \sqrt{2/v} \times \frac{\Gamma(v/2)}{\Gamma((v-1)/2)}$.

2.2. Bias in estimating C_{pk} with $\bar{\bar{X}} - R$ chart information

From Equation (10), $E[\hat{C}_{pk}^{(R)}]$ can be alternatively written as

$$E[\hat{C}_{pk}^{(R)}] = \{f^{(R)}(m, n)\} \times \frac{d}{3\sigma} \times \left(1 - \frac{\mu_f^*}{d} \right), \quad (11)$$

where $f^{(R)}(m, n) = d_2/(b_v \times d_2^*)$, which is a function of m and n only.

Thus, from Equations (3) and (11), the expression for the bias, in estimating C_{pk} using $\hat{C}_{pk}^{(R)}$, is

$$\begin{aligned} \text{bias}_{C_{pk}}(\hat{C}_{pk}^{(R)}) &= E[\hat{C}_{pk}^{(R)}] - C_{pk} \\ &= \frac{d}{3\sigma} \times \{f^{(R)}(m, n) - 1\} + \frac{1}{3} \times \left[\frac{|\mu - M|}{\sigma} - \{f^{(R)}(m, n)\} \times \frac{\mu_f^*}{\sigma} \right] \\ &= \frac{d}{3\sigma} \times \text{bias}_1^{(R)} + \frac{1}{3} \times \text{bias}_2^{(R)}, \end{aligned} \quad (12)$$

where $\text{bias}_1^{(R)} = f^{(R)}(m, n) - 1$ and $\text{bias}_2^{(R)} = |\mu - M|/\sigma - \{f^{(R)}(m, n)\} \times \mu_f^*/\sigma$.

From Equation (12), it is evident that $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(R)}) \rightarrow 0$, if and only if $f^{(R)}(m, n) \rightarrow 1$ and $(\mu_f^*/\sigma - |\mu - M|/\sigma) \rightarrow 0$ simultaneously.
Let

$$\mu_f^*/\sigma = \begin{cases} \left(\frac{\mu_f^*}{\sigma}\right)_+ & \text{for } \mu > M, \\ \left(\frac{\mu_f^*}{\sigma}\right)_- & \text{for } \mu < M. \end{cases}$$

We know that for $x > 0$, $1 - 2\Phi(-x) = -[1 - 2\Phi(x)]$. Hence, from Equation (6)

$$\frac{\mu_f^*}{\sigma} = \sqrt{\frac{2}{N\pi}} \times e^{-\frac{N(\mu-M)^2}{2\sigma^2}} - \left(\frac{\mu-M}{\sigma}\right) \times \left[1 - 2\Phi\left(\frac{(\mu-M)\sqrt{N}}{\sigma}\right)\right]. \quad (13)$$

Now, for $\mu > M$

$$\left(\frac{\mu_f^*}{\sigma}\right)_+ = \sqrt{\frac{2}{N\pi}} \times e^{-\frac{N(\mu-M)^2}{2\sigma^2}} - \left(\frac{\mu-M}{\sigma}\right) \times \left[1 - 2\Phi\left(\frac{(\mu-M)\sqrt{N}}{\sigma}\right)\right]. \quad (14)$$

Also, for $\mu < M$

$$\begin{aligned} \left(\frac{\mu_f^*}{\sigma}\right)_- &= \sqrt{\frac{2}{N\pi}} \times e^{-\frac{N(\mu-M)^2}{2\sigma^2}} - \left\{-\left(\frac{\mu-M}{\sigma}\right)\right\} \times \left[1 - 2\Phi\left(-\frac{(\mu-M)\sqrt{N}}{\sigma}\right)\right] \\ &= \sqrt{\frac{2}{N\pi}} \times e^{-\frac{N(\mu-M)^2}{2\sigma^2}} - \left\{-\left(\frac{\mu-M}{\sigma}\right)\right\} \times \left[-\left\{1 - 2\Phi\left(\frac{(\mu-M)\sqrt{N}}{\sigma}\right)\right\}\right] \\ &= \sqrt{\frac{2}{N\pi}} \times e^{-\frac{N(\mu-M)^2}{2\sigma^2}} - \left(\frac{\mu-M}{\sigma}\right) \times \left[1 - 2\Phi\left(\frac{(\mu-M)\sqrt{N}}{\sigma}\right)\right]. \end{aligned} \quad (15)$$

Thus, $(\mu_f^*/\sigma)_+ = (\mu_f^*/\sigma)_-$. Therefore, the expression for μ_f^*/σ does not depend upon whether $\mu > M$ or $\mu < M$.

Now, to have a deeper look into the characteristics of $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(R)})$, we consider $|\mu - M|/\sigma = 0(0.5)2.0$ (similar to [38]), $m = 5(1)10, 15, 20, 25$, $n = 2(2)10$ and $N = 10, 30, 50, 70, 100, 150, 200$. The corresponding values of $f^{(R)}(m, n)$ (i.e. $\text{bias}_1^{(R)} + 1$) and $\text{bias}_2^{(R)}$ are given in Tables 1 and 2, respectively. In Table 2, $f^{(R)}(m, n)$ is abbreviated by f .

Following observations can be made from Tables 1 and 2:

- (1) $f^{(R)}(m, n)$ decreases with the increase in at least one of m and n . Hence, increase in sample size and/or number of rational subgroups, will decrease the value of $f^{(R)}(m, n)$, which in turn will decrease $\text{bias}_1^{(R)}$.
However, increasing m and n unboundedly may not always be practically feasible, due to the following reasons:
 - (a) High values of m may introduce heterogeneity among the characteristics of the sampled items, as there is supposed to be substantially large time span between the initial and the final rational subgroup.
 - (b) High values of n , say $n \geq 10$ solicits use of S-chart instead of R-chart (refer [22]), since range may fail to capture the true dispersion scenario for such samples.
- (2) For $5 \leq n < 10$ and $8 \leq m \leq 20$, $f^{(R)}(m, n)$, is fairly close to 1 without extra-ordinary increase in the values of m and/or n .

Table 1. $f^{(R)}(m, n)$ values corresponding to different values of m and n .

m	n								
	2	3	4	5	6	7	8	9	10
5	1.1267	1.0653	1.0391	1.0317	1.0228	1.0193	1.0175	1.0153	1.0135
6	1.1365	1.0515	1.0314	1.0269	1.0194	1.0622	1.0145	1.0125	1.0114
7	1.1069	1.0424	1.0284	1.0218	1.0161	1.014	1.0123	1.0106	1.0099
8	1.0877	1.0361	1.0242	1.0194	1.0144	1.0123	1.0107	1.0095	1.0085
9	1.0742	1.0315	1.0210	1.0175	1.0130	1.0107	1.0095	1.0084	1.0076
10	1.0642	1.0301	1.0186	1.0152	1.0114	1.0096	1.0083	1.0075	1.0068
15	1.0430	1.0185	1.0126	1.0093	1.0076	1.0064	1.0056	1.0049	1.0045
20	1.0296	1.0139	1.0093	1.0069	1.0056	1.0048	1.0042	1.0037	1.0034
25	1.0241	1.0111	1.0074	1.0056	1.0045	1.0038	1.0033	1.0030	1.0027

Table 2. $\text{bias}_2^{(R)}$ values corresponding to different values of $|\mu - M|/\sigma$ and N .

N	$\frac{ \mu - M }{\sigma}$				
	0	0.5	1.0	1.5	2.0
10	-0.2525 f	0.5-0.5154 f	1-1.0001 f	1.5(1 - f)	2(1 - f)
30	-0.1457 f	0.5-0.5003 f	1- f	1.5(1 - f)	2(1 - f)
50	-0.1128 f	0.5-0.50001 f	1- f	1.5(1 - f)	2(1 - f)
70	-0.0954 f	0.5(1 - f)	1- f	1.5(1 - f)	2(1 - f)
100	-0.0798 f	0.5(1 - f)	1- f	1.5(1 - f)	2(1 - f)
150	-0.0651 f	0.5(1 - f)	1- f	1.5(1 - f)	2(1 - f)
200	-0.0564 f	0.5(1 - f)	1- f	1.5(1 - f)	2(1 - f)

- (3) According to Table 2, for $|\mu - M|/\sigma \geq 1.0$, $\text{bias}_2^{(R)} \rightarrow |\mu - M|/\sigma \times (1 - f)$ even for the value of N as small as 10. Also, for $0.5 \leq |\mu - M|/\sigma < 1.0$, $\text{bias}_2^{(R)} \rightarrow |\mu - M|/\sigma \times (1 - f)$ for $N > 50$. Thus, for $f^{(R)}(m, n) \rightarrow 1$, $\text{bias}_2^{(R)} \rightarrow 0$, under these situations. In other words, for sufficiently large N and $|\mu - M|/\sigma \geq 0.5$, $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(R)}) \rightarrow 0$ if and only if $f^{(R)}(m, n) \rightarrow 1$.
- (4) Interestingly, the desired values of m and n in Table 1, also satisfy the desired values of N obtained from Table 2, for $|\mu - M|/\sigma \geq 0.5$.
- (5) For $|\mu - M|/\sigma = 0$, $\text{bias}_2^{(R)}$ reduces very slowly even for $N \geq 200$. This indicates that the performance of $C_{pk}^{(R)}$, as an estimator of C_{pk} is not satisfactory when μ is very close to M in σ unit. However, this is not a very vital drawback of $C_{pk}^{(R)}$, as when $\mu \simeq M$, indicating that the process is perfectly centred with respect to the USL and LSL, C_p is used, instead of C_{pk} , to assess the capability of the process.
- (6) By definition, $d/3\sigma > 0$ and from Table 1 $f^{(R)}(m, n) > 1$. Hence, from Equation (12), even if $\text{bias}_2^{(R)} \rightarrow 0$ and $\text{bias}_1^{(R)}$ is small but not sufficiently close to 0, then $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(R)}) > 0$. Therefore, in general, $C_{pk}^{(R)}$ has a tendency to generate positive bias (however small).

2.3. Plug-in estimator of C_{pk} based on $\bar{X} - S$ chart information and the corresponding expectations

Although, due to its simplistic definition, range is the most popular measure of dispersion in the context of statistical quality control, many a times, the sample size is sufficiently large, say $n > 10$ and/or it varies over subgroups. In such situations, standard deviation out-performs range as a dispersion measure (refer [22]) and hence, the use of $\bar{X} - S$ chart is solicited, rather than $\bar{X} - R$ chart, to check and establish statistical stability of a process.

Proceeding similar to the definition of $\hat{C}_{pk}^{(R)}$ [vide equation (4)] and replacing μ and σ , respectively, by $\bar{\bar{X}}$ and $\bar{S}/c_4$ in Equation (3), the plug-in (natural) estimator of C_{pk} , based on $\bar{X} - S$ chart information, can be defined as

$$\hat{C}_{pk}^{(S)} = \frac{dc_4}{3} \left\{ 1 - \frac{|\bar{\bar{X}} - M|}{d} \right\} \times \frac{1}{\bar{S}}. \quad (16)$$

Here, c_4 is a function of the sample size n (refer [22]).

Following Chatterjee and Chakraborty [23], $\bar{S}^2/\sigma^2 \sim 1/m(N-m) \times \chi_{m(N-m)}^2$. Hence,

$$\begin{aligned} E[\hat{C}_{pk}^{(S)}] &= \frac{dc_4}{3} \left\{ 1 - E \left(\frac{|\bar{\bar{X}} - M|}{d} \right) \right\} \times E(\bar{S}^{-1}) \\ &= \frac{dc_4}{3\sigma b_{m(N-m)}} \times \left(1 - \frac{\mu_f^*}{d} \right), \end{aligned} \quad (17)$$

where $b_{m(N-m)} = \sqrt{2/m(N-m)} \times \Gamma(m(N-m)/2)/\Gamma(m(N-m)-1/2)$.

2.4. Bias in estimating C_{pk} with $\bar{X} - S$ chart information

From Equation (17), $E[\hat{C}_{pk}^{(S)}]$ can be alternatively written as

$$E[\hat{C}_{pk}^{(S)}] = \{f^{(S)}(m, n)\} \times \frac{d}{3\sigma} \times \left(1 - \frac{\mu_f^*}{d} \right), \quad (18)$$

where $f^{(S)}(m, n) = c_4/b_{m(N-m)}$, which solely is a function of m and n .

Thus, from Equations (3) and (18), the expression for the bias, in estimating C_{pk} using $\hat{C}_{pk}^{(S)}$, is

$$\begin{aligned} \text{bias}_{C_{pk}}(\hat{C}_{pk}^{(S)}) &= E[\hat{C}_{pk}^{(S)}] - C_{pk} \\ &= \frac{d}{3\sigma} \times \{f^{(S)}(m, n) - 1\} + \frac{1}{3} \times \left[\frac{|\mu - M|}{\sigma} - \{f^{(S)}(m, n)\} \times \frac{\mu_f^*}{\sigma} \right] \\ &= \frac{d}{3\sigma} \times \text{bias}_1^{(S)} + \frac{1}{3} \times \text{bias}_2^{(S)}, \end{aligned} \quad (19)$$

where $\text{bias}_1^{(S)} = f^{(S)}(m, n) - 1$ and $\text{bias}_2^{(S)} = |\mu - M|/\sigma - \{f^{(S)}(m, n)\} \times \mu_f^*/\sigma$.

Thus, similar to the case of $\hat{C}_{pk}^{(R)}$, in Section 2.2, here also it is evident that, $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(S)}) \rightarrow 0$, if and only if $f^{(S)}(m, n) \rightarrow 1$ and $|\mu - M|/\sigma \rightarrow \mu_f^*/\sigma$ simultaneously.

While studying the properties of $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(S)})$, it is interesting to note that, Table 2 will remain valid for $\text{bias}_2^{(S)}$ as well, with the only difference being that, here f stands for $f^{(S)}(m, n)$ instead of $f^{(R)}(m, n)$. Hence, it is sufficient to compute the values of $f^{(S)}(m, n)$ and consequently, those of $\text{bias}_1^{(S)}$ for various values of m and n . For this, we consider $m = 5, 6, 7$ and $n = 10(1)15$. The corresponding values of $f^{(S)}(m, n)$ (i.e. $\text{bias}_1^{(S)} + 1$) is given in Table 3.

From the above discussion and from Table 3, following observations can be made:

- (1) For $m \geq 7$, $f^{(S)}(m, n) \rightarrow 1$ for any sample size $n (\geq 10)$.

Table 3. $f^{(S)}(m, n)$ values corresponding to different values of m and n .

m	n					
	10	11	12	13	14	15
5	0.9760	0.9783	0.9803	0.9819	0.9837	1.0
6	0.9750	1.0	1.0	1.0	1.0	1.0
7	1.0	1.0	1.0	1.0	1.0	1.0

- (2) For $m=6$, n should be at least 11; while for $m \leq 5$, n should be at least 15. However in practice, $m \leq 5$ is hardly observed.
- (3) $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(S)})$ tends to 0 more rapidly than $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(R)})$. This indicates towards the better performance of $\hat{C}_{pk}^{(S)}$ as compared to $\hat{C}_{pk}^{(R)}$ in estimating C_{pk} .
- (4) Unlike $f^{(R)}(m, n)$, $f^{(S)}(m, n) \leq 1$.
- (5) As can be observed from Table 2, which is valid for $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(S)})$ as well, for $0.5 \leq |\mu - M|/\sigma \leq 2.0$ and $N \geq 50$, $\text{bias}_2^{(S)} \rightarrow (1 - f) \times |\mu - M|/\sigma$. Here, $N \geq 50$ satisfy the desirable values of m and n as obtained from Table 3.
- (6) By definition, $d/3\sigma > 0$ and from Table 3, $f^{(S)}(m, n) \leq 1$. Hence, from Equation (19), even if $\text{bias}_2^{(S)} \rightarrow 0$ and $\text{bias}_1^{(S)}$ is small but not sufficiently close to 0, as may sometimes be the case with $m \leq 5$, then $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(S)}) < 0$. Therefore, in general, $C_{pk}^{(S)}$ has a tendency to generate negative bias (however small).

3. Process capability control charts of C_{pk}

We shall now design the control limits of the process capability control charts of C_{pk} based on information from the corresponding $\bar{X}$ –R and $\bar{X}$ –S charts following the approach suggested by Chatterjee and Chakraborty [23].

3.1. Process capability control chart of C_{pk} using $\bar{X}$ –R chart information

From Equation (4), $\hat{C}_{pk}^{(R)} = (dd_2/3)\{1 - |\bar{X} - M|/d\} \times 1/\bar{R}$. Also, following [37], $\bar{R} \sim (\sigma d_2^*/\sqrt{v}) \times \chi_v$ and $|\bar{X} - M| \sim FN(\mu_f^*, \sigma_f^{*2})$, where, μ_f^* and σ_f^{*2} are defined in Equations (6) and (7), respectively. Now, let $\tau_\alpha^{(FN)}(\mu_f^*, \sigma_f^{*2})$ denotes the upper $\alpha\%$ –point of a folded normal distribution with mean μ_f^* and variance σ_f^{*2} . Thus,

$$\begin{aligned}
 1 - \alpha &= P \left[\left(\frac{\sigma d_2^*}{\sqrt{v}} \right) \times \chi_{1-\frac{\alpha}{2}, v} \leq \bar{R} \leq \left(\frac{\sigma d_2^*}{\sqrt{v}} \right) \times \chi_{\frac{\alpha}{2}, v} \right] \\
 &= P \left[\frac{dd_2}{3} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \frac{\sqrt{v}}{\sigma d_2^*} \times \chi_{\frac{\alpha}{2}, v}^{-1} \leq \frac{dd_2}{3} \times \left\{ 1 - \frac{|\bar{X} - M|}{d} \right\} \times \frac{1}{\bar{R}} \right. \\
 &\leq \left. \frac{dd_2}{3} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \frac{\sqrt{v}}{\sigma d_2^*} \times \chi_{1-\frac{\alpha}{2}, v}^{-1} \right] \\
 &= P \left[\frac{dd_2}{3} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \frac{\sqrt{v}}{\sigma d_2^*} \times \chi_{\frac{\alpha}{2}, v}^{-1} \leq \hat{C}_{pk}^{(R)} \right. \\
 &\leq \left. \frac{dd_2}{3} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \frac{\sqrt{v}}{\sigma d_2^*} \times \chi_{1-\frac{\alpha}{2}, v}^{-1} \right] \text{ from Equation (4).} \quad (20)
 \end{aligned}$$

Hence, from Equation (20), the upper control limit (UCL), central line (CL) and lower control limit (LCL) of C_{pk} control chart, based on information from the corresponding $\bar{X}$ –R chart, are, respectively, as follows:

$$\begin{aligned} \text{UCL}_{C_{pk}^{(R)}} &= \frac{dd_2}{3} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \frac{\sqrt{v}}{\sigma d_2^*} \times \chi_{1-\frac{\alpha}{2}, v}^{-1}, \\ \text{CL}_{C_{pk}^{(R)}} &= C_{pk}, \\ \text{LCL}_{C_{pk}^{(R)}} &= \frac{dd_2}{3} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \frac{\sqrt{v}}{\sigma d_2^*} \times \chi_{\frac{\alpha}{2}, v}^{-1}. \end{aligned} \quad (21)$$

Note that, since the control limits in Equation (21) involve unknown parameters, namely μ and σ , which are unobservable in practice, they are to be replaced by $\bar{\bar{X}}$ and $\bar{R}/d_2$, respectively. Accordingly, the modified control limits of C_{pk} control chart, based on the information from the corresponding $\bar{X}$ –R chart, will be

$$\begin{aligned} \text{UCL}_{\hat{C}_{pk}^{(R)}} &= \frac{dd_2^2}{3d_2^*} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)}(\hat{\mu}_f^{*(R)}, \hat{\sigma}_f^{*(R)2}) \right\} \times \frac{\sqrt{v}}{\bar{R}} \times \chi_{1-\frac{\alpha}{2}, v}^{-1}, \\ \text{CL}_{\hat{C}_{pk}^{(R)}} &= \hat{C}_{pk}^{(R)}, \\ \text{LCL}_{\hat{C}_{pk}^{(R)}} &= \frac{dd_2^2}{3d_2^*} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)}(\hat{\mu}_f^{*(R)}, \hat{\sigma}_f^{*(R)2}) \right\} \times \frac{\sqrt{v}}{\bar{R}} \times \chi_{\frac{\alpha}{2}, v}^{-1}, \end{aligned} \quad (22)$$

where $\hat{\mu}_f^{*(R)}$ and $\hat{\sigma}_f^{*(R)2}$ are obtained by replacing μ by $\bar{\bar{X}}$ and σ by $\bar{R}/d_2$ in Equations (6) and (7), respectively.

In this context, Chatterjee and Chakraborty [39] have provided a simple algorithm to compute the values of $\tau_{\alpha}^{(FN)}(\cdot)$.

3.2. Process capability control chart of C_{pk} using $\bar{X}$ –S chart information

From Equation (16), $\hat{C}_{pk}^{(S)} = (dc_4/3)\{1 - |\bar{\bar{X}} - M|/d\} \times 1/\bar{S}$. Also, $\bar{S}^2/\sigma^2 \sim 1/m(N-m)\chi_{m(N-m)}^2$ (refer [23]), that is, $\bar{S} \sim \sigma/\sqrt{m(N-m)} \times \chi_{m(N-m)}$. Hence,

$$\begin{aligned} 1 - \alpha &= P \left[\frac{\sigma}{\sqrt{m(N-m)}} \times \chi_{1-\frac{\alpha}{2}, m(N-m)} \leq \bar{S} \leq \frac{\sigma}{\sqrt{m(N-m)}} \times \chi_{\frac{\alpha}{2}, m(N-m)} \right] \\ &= P \left[\frac{\sqrt{m(N-m)}}{\sigma} \times \frac{dc_4}{3} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \chi_{\frac{\alpha}{2}, m(N-m)}^{-1} \right. \\ &\leq \frac{dc_4}{3} \left\{ 1 - \frac{|\bar{\bar{X}} - M|}{d} \right\} \times \frac{1}{\bar{S}} \leq \frac{\sqrt{m(N-m)}}{\sigma} \times \frac{dc_4}{3} \\ &\quad \left. \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \times \chi_{1-\frac{\alpha}{2}, m(N-m)}^{-1} \right] \end{aligned}$$

$$\begin{aligned}
&= P \left[\frac{dc_4 \sqrt{m(N-m)}}{3\sigma} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)} \left(\mu_f^*, \sigma_f^{*2} \right) \right\} \times \chi_{\frac{\alpha}{2}, m(N-m)}^{-1} \leq \hat{C}_{pk}^{(S)} \right. \\
&\leq \left. \frac{dc_4 \sqrt{m(N-m)}}{3\sigma} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)} \left(\mu_f^*, \sigma_f^{*2} \right) \right\} \times \chi_{1-\frac{\alpha}{2}, m(N-m)}^{-1} \right] \quad \text{from Equation (16)}.
\end{aligned} \tag{23}$$

Therefore, from Equation (23) and proceeding similar to Section 3.1, the control limits of C_{pk} control chart, based on the information from the corresponding $\bar{X}-S$ chart, will be

$$\begin{aligned}
UCL_{\hat{C}_{pk}^{(S)}} &= \frac{dc_4^2 \sqrt{m(N-m)}}{3\bar{S}} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)} (\hat{\mu}_f^{*(S)}, \hat{\sigma}_f^{*(S)2}) \right\} \times \chi_{1-\frac{\alpha}{2}, m(N-m)}^{-1}, \\
CL_{\hat{C}_{pk}^{(S)}} &= \hat{C}_{pk}^{(S)}, \\
LCL_{\hat{C}_{pk}^{(S)}} &= \frac{dc_4^2 \sqrt{m(N-m)}}{3\bar{S}} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)} (\hat{\mu}_f^{*(S)}, \hat{\sigma}_f^{*(S)2}) \right\} \times \chi_{\frac{\alpha}{2}, m(N-m)}^{-1},
\end{aligned} \tag{24}$$

where $\hat{\mu}_f^{*(S)}$ and $\hat{\sigma}_f^{*(S)2}$ are obtained by replacing μ by $\bar{X}$ and σ by $\bar{S}/c_4$ in Equations (6) and (7), respectively.

Thus, the operational procedure for the control chart of C_{pk} using $\bar{X}-R$ and $\bar{X}-S$ chart information can be designed through the following steps:

- Step 1:** Draw random samples of size n from m rational subgroups.
 - Step 2:** Draw $\bar{X}-R$ or $\bar{X}-S$ chart from the data to check whether the process mean and dispersion are under statistical control. The selection between R-chart and S-chart should be based on the sample size. If $n \leq 9$ R-chart should be drawn and if $n \geq 11$, S-chart is preferable. For $n = 10$, either of the two charts can be considered, although the decision should be based on the knowledge of the process and past experiences.
 - Step 3:** If the mean and the dispersion of the process are found to be statistically stable, draw C_{pk} control chart using $\bar{X}-R$ or $\bar{X}-S$ chart, whichever was selected in Step 2.
 - Step 4:** If in the C_{pk} control chart, all the points lie within the UCL and LCL, the process can be considered as consistently capable throughout the entire production cycle and hence its performance can be summarized through a single PCI value (vide equations (4) or (16), whichever is applicable).
 - Step 5:** If at least one point is below the LCL of C_{pk} control chart, the process cannot be considered as consistently capable. Although, any point above the UCL should generally indicate towards extraordinarily good performance of the process at that point of production, this phenomenon should be checked for possible assignable cause or computational error or any such undesirable situations, before making such hopeful decision.
- When the process is not consistently capable, its performance should not be summarized through a single C_{pk} or any other PCI value.

3.3. β -Risk and average run length for C_{pk} control chart

β -risk and the corresponding average run length (ARL) are two of the most important performance yardsticks of a control chart.

Let us assume that C_{pk_0} is a known specific value of C_{pk} and k' is the change in C_{pk_0} value. Then, from Equation (21), the β -risk for C_{pk} control chart, based on $\bar{X}-R$ chart information, can be

Table 4. β -Risk and ARL for C_{pk} control chart based on $\bar{X} - R$ chart information.

k'	C_{pk_0}				
	1.00	1.20	1.40	1.60	1.80
0.1	0.1807(2)	0.2209(2)	0.2610(2)	0.3012(2)	0.3413(2)
0.2	0.1606(2)	0.2008(2)	0.2409(2)	0.2811(2)	0.3213(2)
0.3	0.1405(2)	0.1807(2)	0.2209(2)	0.2610(2)	0.3012(2)
0.4	0.1205(2)	0.1606(2)	0.2008(2)	0.2409(2)	0.2811(2)
0.5	0.1000(2)	0.1405(2)	0.1807(2)	0.2209(2)	0.2610(2)
0.6	0.0803(2)	0.1205(2)	0.1606(2)	0.2008(2)	0.2409(2)
0.7	0.0602(2)	0.1003(2)	0.1405(2)	0.1807(2)	0.2209(2)
0.8	0.0401(2)	0.0803(2)	0.1205(2)	0.1606(2)	0.2008(2)
0.9	0.0200(2)	0.0602(2)	0.1004(2)	0.1405(2)	0.1807(2)
1.00	0(1)	0.0401(2)	0.0803(2)	0.1205(2)	0.1606(2)

derived as

$$\begin{aligned}
 \beta &= P[\text{LCL}_{C_{pk}}^{(R)} \leq \hat{C}_{pk}^{(R)} \leq \text{UCL}_{C_{pk}}^{(R)} \mid C_{pk} = C_{pk_0} - k'], \quad \text{since } C_{pk} \text{ is a higher the better type index} \\
 &= P \left[\hat{C}_{pk}^{(R)} \leq \frac{d_2 \sqrt{v}}{d_2^* \left\{ 1 - \frac{|\mu - M|}{d} \right\}} \times C_{pk} \times \chi_{1-\alpha/2, v}^{-1} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \mid C_{pk} = C_{pk_0} - k' \right] \\
 &\quad - P \left[\hat{C}_{pk}^{(R)} \leq \frac{d_2 \sqrt{v}}{d_2^* \left\{ 1 - \frac{|\mu - M|}{d} \right\}} \times C_{pk} \times \chi_{\alpha/2, v}^{-1} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \mid C_{pk} = C_{pk_0} - k' \right]. \quad (25)
 \end{aligned}$$

Similarly, from Equation (23), the β -risk for C_{pk} control chart, based on $\bar{X} - S$ chart information, can be derived as

$$\begin{aligned}
 \beta &= P[\text{LCL}_{C_{pk}}^{(S)} \leq \hat{C}_{pk}^{(S)} \leq \text{UCL}_{C_{pk}}^{(S)} \mid C_{pk} = C_{pk_0} - k'] \\
 &= P \left[\hat{C}_{pk}^{(S)} \leq \frac{a \times c_4}{1 - \frac{|\mu - M|}{d}} \times C_{pk} \times \chi_{1-\alpha/2, a^2}^{-1} \times \left\{ 1 - \frac{1}{d} \times \tau_{1-\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \mid C_{pk} = C_{pk_0} - k' \right] \\
 &\quad - P \left[\hat{C}_{pk}^{(S)} \leq \frac{a \times c_4}{1 - \frac{|\mu - M|}{d}} \times C_{pk} \times \chi_{\alpha/2, a^2}^{-1} \times \left\{ 1 - \frac{1}{d} \times \tau_{\frac{\alpha}{2}}^{(FN)}(\mu_f^*, \sigma_f^{*2}) \right\} \mid C_{pk} = C_{pk_0} - k' \right], \quad (26)
 \end{aligned}$$

where $a = \sqrt{m(N - m)}$.

Thus, β -risk for C_{pk} chart depends upon the values of USL, LSL, μ , σ , m and n .

For USL = 3.8 and LSL = 2.0, $\mu = 3$ and $\sigma = 0.2$, $n = 5$ and $m = 5$, Tables 4 and 5 give the β and ARL (within parentheses) values of C_{pk} control chart corresponding to different magnitudes of shift in C_{pk} values using $\bar{X} - R$ chart and $\bar{X} - S$ chart information, respectively.

From Tables 4 and 5, following can be observed:

- (1) For C_{pk} control chart using both $\bar{X} - R$ chart and $\bar{X} - S$ chart information, the β -risk increases with the increase in C_{pk_0} value, for fixed k' .
- (2) For C_{pk} control chart using both $\bar{X} - R$ chart and $\bar{X} - S$ chart information, the β -risk decreases with the increase in k' value, for a fixed C_{pk_0} value.
- (3) For C_{pk} control chart using $\bar{X} - R$ chart information, ARL is constant, namely 2, irrespective of the changes in C_{pk_0} or k' values, except for $C_{pk_0} = 1.00$ and $k' = 1.0$, for which, ARL is 1.
- (4) From Table 4, all the β values are considerably small and consequently, the values of power are considerably high, which is highly desirable. The low values of ARL are also indicating towards

Table 5. β -Risk and ARL for C_{pk} control chart based on $\bar{X} - S$ chart information.

k'	C_{pk_0}				
	1.00	1.20	1.40	1.60	1.80
0.1	0.4890(2)	0.5977(3)	0.7063(4)	0.8150(6)	0.9237(14)
0.2	0.4346(2)	0.5433(3)	0.6520(3)	0.7607(5)	0.8693(8)
0.3	0.3803(2)	0.4889(2)	0.5977(3)	0.7063(4)	0.8150(6)
0.4	0.3260(2)	0.4347(2)	0.5433(3)	0.6520(3)	0.7607(5)
0.5	0.2717(2)	0.3803(2)	0.4889(2)	0.5977(3)	0.7063(4)
0.6	0.2173(2)	0.3260(2)	0.4347(2)	0.5433(3)	0.6520(3)
0.7	0.1630(2)	0.2717(2)	0.3803(2)	0.4890(2)	0.5977(2)
0.8	0.1087(2)	0.2173(2)	0.3260(2)	0.4347(2)	0.5433(3)
0.9	0.0543(2)	0.1630(2)	0.2717(2)	0.3803(2)	0.4890(2)
1.00	0(1)	0.1087(2)	0.2173(2)	0.3260(2)	0.4347(2)

the satisfactory performance of C_{pk} control chart using $\bar{X} - R$ chart information for the given values of USL, LSL, μ , σ , m and n , to detect process shift.

- (5) From Table 5, ARL increases with the increase in C_{pk_0} value, for fixed k' and decreases with the increase in k' value, for a fixed C_{pk_0} value.
- (6) For C_{pk} control chart using $\bar{X} - S$ chart information, the β values are not quite small for small values of C_{pk_0} and k' . However, for moderate to high values of k' , say $k' \geq 0.6$ and small values of C_{pk_0} , namely $C_{pk_0} \leq 1.20$, the β values are quite small.
Thus, as can be observed from Table 5, C_{pk} control chart using $\bar{X} - S$ chart information performs better for small C_{pk_0} values and for moderate to large process shifts.
- (7) From Table 4, C_{pk} control chart using $\bar{X} - R$ chart information performs satisfactorily (in terms of β value) for almost all values of k' and for all C_{pk_0} values, except the combinations $C_{pk_0} \geq 1.60$ and $k' = k' \geq 0.7$, that is, when the values of both C_{pk_0} and k' are considerably high.

Note that, although Tables 4 and 5 are generated for specific values of USL, LSL, μ , σ , m and n , the observations which are enlisted above are true for other combinations of these parameters as well. Thus, Tables 4 and 5 are, in some sense, indicative of the nature of β -risk and ARL of C_{pk} control chart.

3.4. Confidence intervals of the plug-in estimators of C_{pk} based on the control chart information

We have already discussed about the control limits of the plug-in estimators of C_{pk} , based on $\bar{X} - R$ and $\bar{X} - S$ chart information. Since these control limits are, by definition, probability limits at α level of significance, these can be used as confidence intervals of C_{pk} as well.

Thus, for $\hat{C}_{pk}^{(R)}$, the $100(1 - \alpha)\%$ confidence interval will be $[LCL_{\hat{C}_{pk}^{(R)}}, UCL_{\hat{C}_{pk}^{(R)}}]$, when $\bar{X} - R$ chart information is used. Similarly, for $\hat{C}_{pk}^{(S)}$, the $100(1 - \alpha)\%$ confidence interval will be $[LCL_{\hat{C}_{pk}^{(S)}}, UCL_{\hat{C}_{pk}^{(S)}}]$, when $\bar{X} - S$ chart information is used.

3.5. Comparison of the proposed C_{pk} control charts with the existing ones

We are now in a position to compare the performance of our newly proposed C_{pk} control charts with the existing ones.

4. A simulation study

We shall now carry out a simulation study to strengthen the theory developed in this article.

Let us consider a hypothetical process for which the underlying distribution of the concerned quality characteristic is normal with $\mu = 3$ and $\sigma = 0.2$. Suppose from each of 10 rational subgroups of the said process, a random sample of size 5 is drawn. Thus, here $m = 10$ and $n = 5$ and hence $N = 50$. Therefore, $\bar{X} - R$ chart will be appropriate here to check the stability of the process. From the $\bar{X}$ chart and R chart of this simulated data (vide Figures 1 and 2, respectively), it is observed that the process is statistically under control.

Suppose for the concerned quality characteristic, $USL = 3.8$ and $LSL = 2.0$. Thus, $d = 0.9$ and $M = 2.9$. For $n = 5$, $d_2 = 2.326$ and $d_3 = 0.864$. Hence, $d_2^* = 2.3398$ and $v = 36.4861 \simeq 36$. From

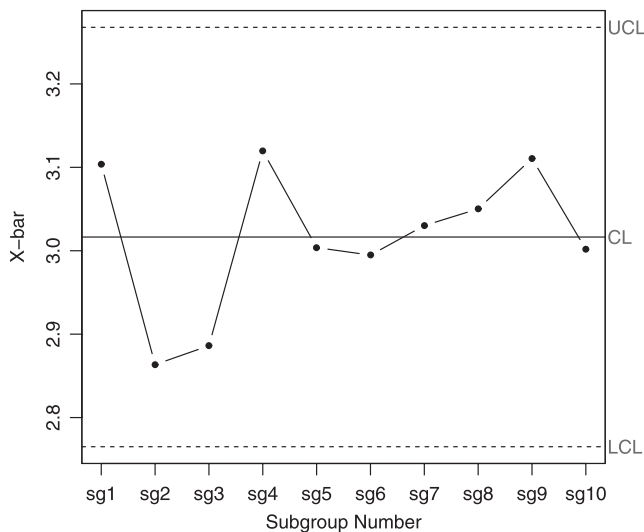


Figure 1. $\bar{X}$ chart for the data of Example 3.

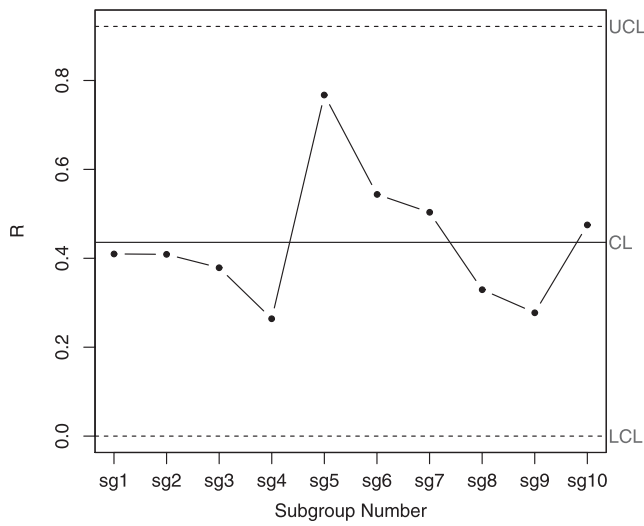


Figure 2. R chart for the data of Example 3.

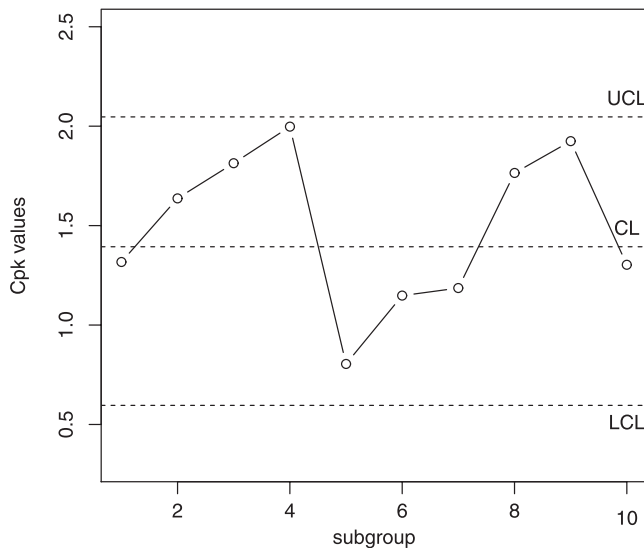


Figure 3. C_{pk} control chart using $\bar{X}$ – R chart information based on data of Section 4.

the data, $\bar{\bar{X}} = 3.0164$, $\bar{R} = 0.4359$ and hence $\hat{\sigma} = 0.1874$. Also, $\hat{\mu}_f^{*(R)} = 0.1164$ and $\hat{\sigma}_f^{*(R)2} = 0.0007$. Hence, $\hat{C}_{pk}^{(R)} = 1.3937$.

We shall now investigate whether it is adequate to summarize the process capability merely based on a single valued estimated PCI like $\hat{C}_{pk}^{(R)}$. For this, we need to use process capability control chart of C_{pk} . Following Chatterjee and Chakraborty [39], $\tau_{0.975}^{(FN)} = 0.00713$ and $\tau_{0.025}^{(FN)} = 0.486$. Hence, for $\alpha = 0.05$, that is, at 5% level of significance, from Equation (22) we have, $UCL_{\hat{C}_{pk}^{(R)}} = 2.0467$, $CL_{\hat{C}_{pk}^{(R)}} = 1.3937$ and $LCL_{\hat{C}_{pk}^{(R)}} = 0.5960$.

The corresponding C_{pk} control chart is given in Figure 3.

From Figure 3, it is easy to observe that, none of the estimated C_{pk} values for the individual subgroups lie beyond $LCL_{\hat{C}_{pk}^{(R)}}$ or $UCL_{\hat{C}_{pk}^{(R)}}$. Hence, apart from being statistically stable, the process can also be expected to perform consistently. Therefore its capability can truly be summarized through $\hat{C}_{pk}^{(R)} = 1.3937$, which indicates towards the satisfactory performance of the process.

Let us now consider the same process with, that is, with $USL = 3.8$ and $LSL = 2.0$. Table 6 contains the UCL, CL and LCL values for various combinations of μ , σ , n and m .

Note that, unlike the simulated example, the values of the UCL CL and LCL are computed in Table 6 assuming that the actual values of μ and σ are known and hence the UCL, CL and LCL values in the simulated example and those in the table differ for $n = 5$ and $m = 10$.

The following observations can be made from Table 6:

- (1) For fixed n , value of UCL decreases and LCL increases with the increase in m . Consequently, the span between UCL and LCL decreases making the control limits more stringent. The same thing is true for fixed m and varying n as well.
- (2) Since, $CL = C_{pk}$ and this is a function of μ , σ , USL and LSL , provided μ and σ values are known, the values of CL remain fixed for each set of (μ, σ) , whatever be the value of m and n .
- (3) For fixed σ , if μ decreases, the UCL and LCL values decrease with the increase in m and n .
- (4) For fixed μ , if σ decreases, the UCL and LCL values decrease with the increase in m and n .

Table 6. Table showing values of UCL, CL and LCL for various combinations of μ , σ , n and m .

n	m	UCL	CL	LCL
For $\mu = 3, \sigma = 0.2$				
3	5	4.5371	1.3333	0.4535
3	10	3.7059	1.3333	0.5027
3	15	3.4178	1.3333	0.5277
3	20	3.2645	1.3333	0.5436
5	5	5.0950	1.3333	0.5027
5	10	4.4873	1.3333	0.5436
5	15	4.2583	1.3333	0.5636
5	20	4.1319	1.3333	0.5763
9	5	5.7897	1.3333	0.5401
9	10	5.3123	1.3333	0.5735
9	15	5.1236	1.3333	0.5895
9	20	5.0170	1.3333	0.5994
For $\mu = 2.6, \sigma = 0.2$				
3	5	4.4754	1.0	0.2340
3	10	3.6556	1.0	0.2595
3	15	3.3714	1.0	0.2724
3	20	3.2201	1.0	0.2806
5	5	5.0258	1.0	0.2595
5	10	4.4263	1.0	0.2806
5	15	4.2005	1.0	0.2909
5	20	4.0758	1.0	0.2974
9	5	5.7110	1.0	0.2788
9	10	5.2402	1.0	0.2960
9	15	5.0540	1.0	0.3042
9	20	4.9488	1.0	0.3094
For $\mu = 3, \sigma = 0.4$				
3	5	2.2537	0.6667	0
3	10	1.8409	0.6667	0
3	15	1.6978	0.6667	0
3	20	1.6216	0.6667	0
5	5	2.5309	0.6667	0
5	10	2.2290	0.6667	0
5	15	2.1153	0.6667	0
5	20	2.0525	0.6667	0
9	5	2.8760	0.6667	0
9	10	2.6388	0.6667	0
9	15	2.5451	0.6667	0
9	20	2.4921	0.6667	0

- (5) In particular, for $\mu = 3$ and $\sigma = 0.4$, with the same USL and LSL as before, the LCL value will always be negative, whatever be the value of m and n and hence we shall consider LCL as zero, since by convention PCIs are non-negative.

5. Applications

We are now in a position to discuss some numerical examples to supplement the theory developed so far in the present article.

5.1. Example 1

In the present example, we consider the data pertaining to a forging process which produces piston rings for an automotive engine. The inside diameter of a forged piston ring is the quality characteristic under consideration. These data have been already discussed by Lin and Sheen [7], while discussing the performance of a Bayesian-like estimator of C_{pk} based on control chart data.

The data set consists of 25 subgroups, each having 5 sample observations. Thus, $m = 25$, $n = 5$ and hence $N = 125$. The USL and LSL of the inside diameters of the forged piston rings are $74.00 \pm$

0.05 mm, that is, $USL = 74.05$ and $LSL = 73.95$. Thus, $d = 0.05$ and $M = 74.0$. Now, following Montgomery [22], for $n = 5$, $d_2 = 2.326$, $d_3 = 0.864$. Hence, $d_2^* = 2.3315$ and $v = 90.8438 \approx 91$. The $\bar{X}$ and R control charts corresponding to the present data are given in Figures 4 and 5.

From these figures, it is evident that, the process can be considered to be statistically in control. From the data, $\bar{X} = 74.0012$, $\bar{R} = 0.02324$ and hence $\hat{\sigma} = 0.01$. Thus, from Equations (6) and (7), $\hat{\mu}_f^{*(R)} = 0.001$ and $\hat{\sigma}_f^{*(R)2} = 0.0000006$. Hence, from the given data, $\hat{C}_{pk}^{(R)} = 1.6289$.

Now, from Table 1, for $m = 25$ and $n = 5$, $f^{(R)}(m, n) = 1.0056$ and from Table 2, for $N = 125$ and $|\bar{X} - M| \times d_2 / \bar{R} = 0.12$, $bias_2^{(R)} = -0.0081$. Hence, from Equation (12), $bias_{C_{pk}}(\hat{C}_{pk}^{(R)}) = 0.0066$, which is quite small.

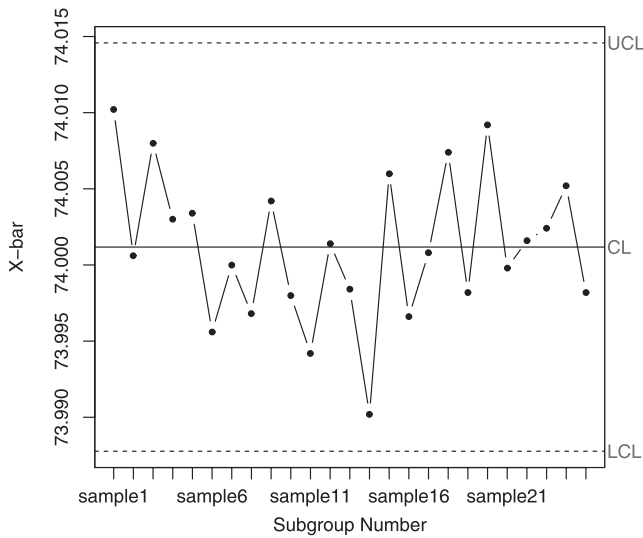


Figure 4. $\bar{X}$ chart for the data of Example 1.

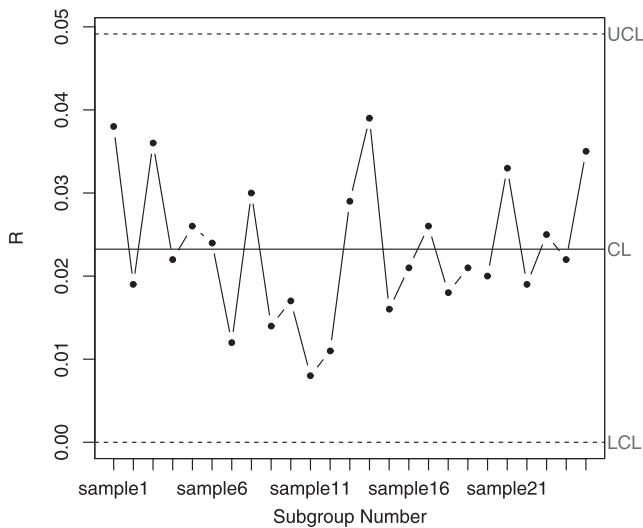


Figure 5. R chart for the data of Example 1.

Interestingly, following Lin and Sheen [7], the estimated value of C_{pk} , namely $\hat{C}_{pk(R)}$ is 1.63, which is very close to $\hat{C}_{pk}^{(R)}$. However, unlike $\hat{C}_{pk(R)}$, we do not have to rely on any prior information for computing $\hat{C}_{pk}^{(R)}$. This strongly argues in favour of classical estimation (using control chart information) of C_{pk} rather than Bayesian or Bayesian-like estimation. This is somewhat similar to the observation made by Pearn et al. [21], who also have advocated the application of classical method to evaluate process capability, rather than using Bayesian approach, in the context of the lower control bound of C_{pk} .

Although, both $\hat{C}_{pk}^{(R)}$ and $\hat{C}_{pk(R)}$ indicate towards the satisfactory performance of the forging process, we shall now investigate whether it is adequate to summarize the process capability merely based on a single valued estimated PCI like $\hat{C}_{pk}^{(R)}$. For this, we need to use process capability control chart of C_{pk} . Now, following Chatterjee and Chakraborty [39], $\tau_{0.975}^{(FN)} = 0.0003154$ and $\tau_{0.025}^{(FN)} = 0.02255$. Hence, at $\alpha = 0.05$ level of significance, from Equation (22) we have, $UCL_{\hat{C}_{pk}^{(R)}} = 1.9347$, $CL_{\hat{C}_{pk}^{(R)}} = 1.6289$ and $LCL_{\hat{C}_{pk}^{(R)}} = 0.7979$. The corresponding C_{pk} control chart is given in Figure 6.

Here, it can be observed that, none of the estimated C_{pk} values for the individual subgroups lie below $LCL_{\hat{C}_{pk}^{(R)}}$, while, for 8 out of 25 subgroups, the estimated C_{pk} values lie above $UCL_{\hat{C}_{pk}^{(R)}}$. Although, PCIs being of higher the better type, exceeding $UCL_{\hat{C}_{pk}^{(R)}}$ cannot be considered as a setback of the process (refer [23,25]); the $\hat{C}_{pk}^{(R)}$ values vary highly from subgroup to subgroup, the maximum being 4.2837 (corresponding to the subgroup no. 11) and the minimum being 0.7992 (corresponding to the subgroup no. 14). These indicate towards the highly volatile performance of the process from sample to sample. In this scenario, summarizing the process capability through a single value like $\hat{C}_{pk}^{(R)}$ will merely average out the ups and downs of the capability of the process over various subgroups. Hence, although $\hat{C}_{pk}^{(R)} = 1.6289$ seems to indicate towards the satisfactory performance of the forging process, in reality the process performance is highly volatile and hence summarizing it through $\hat{C}_{pk}^{(R)}$ is not solicited.

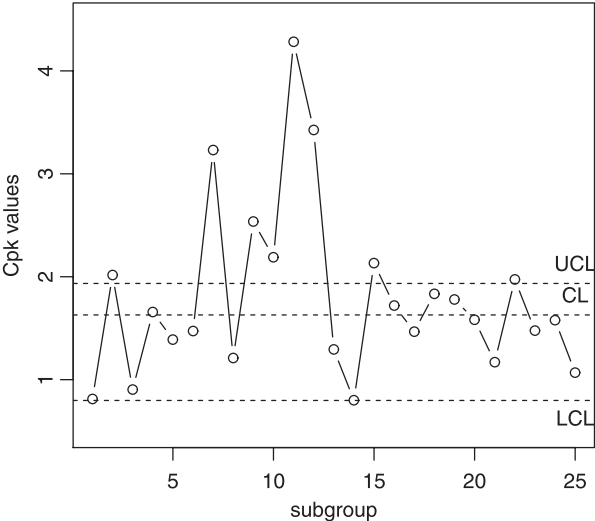


Figure 6. C_{pk} control chart using $\bar{X}$ – R chart information based on data of Example 1.

This also indicates that, despite being under statistical control (as reflected by Figures 4 and 5), the performance of the process is not consistent. Therefore, similar to the observation made by Chatterjee and Chakraborty [23], in the current example, statistical stability does not ensure consistency in the process performance.

5.2. Example 2

In this example, we consider the data which was used earlier by Spiring [25]. Here we have 20 subgroups each having measurements of the concerned quality characteristic for 10 sample units with $USL = 1.2$ and $LSL = 0.8$. Thus, $m = 20$, $n = 10$, $N = 200$, $d = 0.2$, $M = 1.0$. From the $\bar{X}$ chart, R-chart and S-chart for the present data set, it is observed that the process is under statistical control (refer [25]).

Since here $n = 10$, we shall compare the performances of both $\hat{C}_{pk}^{(R)}$ and $\hat{C}_{pk}^{(S)}$ and the corresponding process capability control charts.

5.2.1. Using $\bar{X}$ –R chart information.

For $n = 10$, $d_2 = 3.0775$ and $d_3 = 0.7971$ (refer [22]) and hence, $d_2^* = 3.0816$ and $v = 149.3128 \approx 149$. From the data, $\bar{\bar{X}} = 1.12055$ and $\bar{R} = 0.348$. Hence, $\hat{C}_{pk}^{(R)} = 0.2342$.

Now, from Table 1, for $m = 20$ and $n = 10$, $f^{(R)}(m, n) = 1.0034$ and from Table 2, for $N = 200$ and $|\bar{\bar{X}} - M| \times d_2 / \bar{R} = 1.07$, $\text{bias}_2^{(R)} = -0.0036$. Hence, from Equation (12), $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(R)}) = 0.0008$, which is quite small.

Since $\hat{C}_{pk}^{(R)} = 0.2342$ is quite small as compared to the stipulated threshold value 1 (refer [2]), it is logical to conclude that the process is incapable of producing items within the given specification limits.

We shall now investigate whether it is adequate to summarize the process capability merely based on $\hat{C}_{pk}^{(R)}$. For this, we need to use process capability control chart of C_{pk} . Using Equations (6) and (7), $\hat{\mu}_f^* = 0.12055$ and $\hat{\sigma}_f^{*2} = 0.00006$. Also, following [39], $\tau_{0.975}^{(FN)} = 0.34224$ and $\tau_{0.025}^{(FN)} = 0.00626$. Hence, at $\alpha = 0.05$ level of significance, from Equation (22) we have, $UCL_{\hat{C}_{pk}^{(R)}} = 0.6433$, $CL_{\hat{C}_{pk}^{(R)}} = 0.2342$ and $LCL_{\hat{C}_{pk}^{(R)}} = -0.3761 \approx 0$. The corresponding C_{pk} control chart is given in Figure 7.

Here, it can be observed that, none of the $\hat{C}_{pk}^{(R)}$ values for the individual subgroups lie below $LCL_{\hat{C}_{pk}^{(R)}}$; while only one such value lies above the $UCL_{\hat{C}_{pk}^{(R)}}$. There is also no drastic variation in the values of $\hat{C}_{pk}^{(R)}$ from one subgroup to another, the minimum $\hat{C}_{pk}^{(R)}$ being 0.0405 (corresponding to subgroup no. 12) and the maximum being 0.7829 (corresponding to subgroup no. 19). Therefore, it is logical to expect that the performance of the process is consistent over various subgroups and hence, $\hat{C}_{pk}^{(R)} = 0.2342$ can be aptly considered to represent the estimated process capability.

Interestingly, despite having consistency in performance, the process fails to perform satisfactorily, as has been reflected by the very low value of $\hat{C}_{pk}^{(R)}$. In fact, based on the available data, it has been observed that, under the present process centring and dispersion scenario, the $UCL_{\hat{C}_{pk}^{(R)}}$ is 0.6433, which itself is quite low, as compared to the stipulated threshold value 1 (refer [2]).

5.2.2. Using $\bar{X}$ –S chart information.

For $n = 10$, $c_4 = 0.9727$ (refer [22]). Also, from the data, $\bar{\bar{X}} = 1.12055$ and $\bar{S} = 0.1094$. Hence, $\hat{C}_{pk}^{(S)} = 0.2347$.

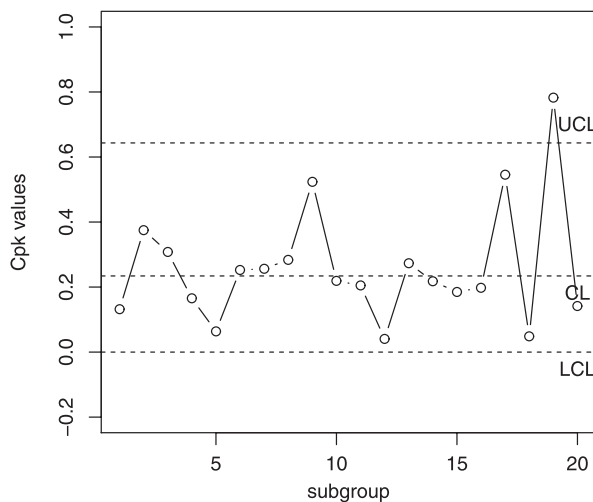


Figure 7. C_{pk} control chart using $\bar{X}$ –R chart information based on data of Example 2.

Now, from Table 3, for $m = 20$ and $n = 10$, $f^{(R)}(m, n) \approx 1.00$ and from Table 2, for $N = 200$ and $|\bar{X} - M| \times d_2/\bar{R} = 1.07$, $\text{bias}_2^{(R)} \approx 0$. Hence, from Equation (19), $\text{bias}_{C_{pk}}(\hat{C}_{pk}^{(S)}) \approx 0$. Therefore, for the present data set, $\hat{C}_{pk}^{(S)}$ performs better than $\hat{C}_{pk}^{(R)}$.

Since $\hat{C}_{pk}^{(S)} = 0.2347$ is quite small as compared to the stipulated threshold value 1 (refer [2]), similar to the case of $\hat{C}_{pk}^{(R)}$, here also it is logical to conclude that the process is incapable of producing items within the given specification limits.

To check the adequacy of $\hat{C}_{pk}^{(S)}$ to summarize the overall process performance, we shall now consider process capability control chart of C_{pk} based on $\bar{X}$ –S chart information. Here, $\tau_{0.975}^{(FN)} = 0.34103$ and $\tau_{0.025}^{(FN)} = 0.00626$ (refer [39]). Hence, from Equation (24), we have, $\text{UCL}_{\hat{C}_{pk}^{(S)}} = 0.5717$, $\text{CL}_{\hat{C}_{pk}^{(S)}} = 0.2347$ and $\text{LCL}_{\hat{C}_{pk}^{(S)}} = -0.3974 \approx 0$. The corresponding C_{pk} control chart is given in Figure 8.

Similar to Figure 7, here also it can be observed that, none of the $\hat{C}_{pk}^{(S)}$ values for the individual subgroups lie below $\text{LCL}_{\hat{C}_{pk}^{(S)}}$; while only one such value lies above the $\text{UCL}_{\hat{C}_{pk}^{(S)}}$. There is also no drastic variation in the values of $\hat{C}_{pk}^{(S)}$ from one subgroup to another. Therefore, it is logical to expect that the performance of the process is consistent over various subgroups and hence, $\hat{C}_{pk}^{(S)} = 0.2347$ can be aptly considered to represent the estimated process capability. Moreover, under the present process centring and dispersion scenario, $\text{UCL}_{\hat{C}_{pk}^{(S)}}$ is 0.5717, which itself is quite low, as compared to the stipulated threshold value 1 (refer [2]).

From both Figures 7 and 8, it is evident that only one estimated C_{pk} value is above the corresponding UCL. This cannot be considered as a setback of a process, as PCIs are in general larger the better type, though proper investigation is required before considering this high value as a benchmark (refer [25]). Also, none of the estimated C_{pk} values lie below the corresponding LCL, which is also indicating towards the consistency of the process performance.

Finally, although, both $\hat{C}_{pk}^{(R)}$, $\hat{C}_{pk}^{(S)}$ and the associated C_{pk} draw similar conclusions regarding the process performance, $\hat{C}_{pk}^{(S)}$ is almost unbiased for C_{pk} , while $\hat{C}_{pk}^{(R)}$ induces slight positive bias. Hence the use of $\hat{C}_{pk}^{(S)}$ is advocated for the present data set. Infact, for $n = 10$, the relative efficiency of R over S,

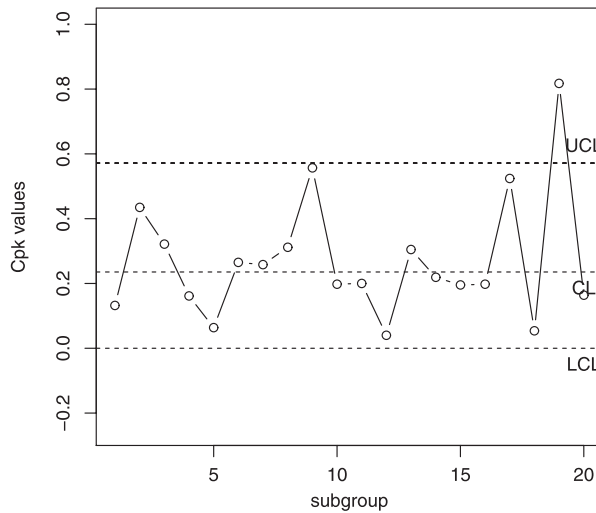


Figure 8. C_{pk} control chart using $\bar{X}$ –S chart information based on data of Example 2.

as a measure of dispersion, is 0.850 and it decreases further as n increases (refer [22]). This also argues in favour of using $\bar{X}$ –S chart information, rather than that of $\bar{X}$ –R chart, in the present scenario.

6. Conclusion

In the present article, we have proposed natural (plug-in) estimators of C_{pk} based on $\bar{X}$ –R chart and $\bar{X}$ –S chart information. We have derived the expressions for the expected values of these estimators and the associated bias. We have also designed the corresponding process capability control charts for the purpose of continual assessment of the process performance. It has been observed that classical estimation of C_{pk} based on $\bar{X}$ –R chart and $\bar{X}$ –S chart information performs better than its Bayesian counterpart. Hence, in practice, for process capability assessment, application of classical approach, which does not require to rely on any historical data, is advocated. Pearn et al. [21] have also made similar observations in the context of lower bound of C_{pk} . It has also been observed that, single valued assessment of process capability is not always adequate to summarize the overall capability of a process and one may have to take refuge to the corresponding process capability control chart.

Two real-life examples and a simulated example have been discussed that reveal the following three different situations which may be encountered in practice:

- (1) In case of the forging process, discussed in Example 1 of Section 5, although the process is statistically stable, the C_{pk} chart reveals that 8 out of 25 subgroup level observed C_{pk} values lie above the UCL, while none of the observed C_{pk} values lie below LCL. Thus, despite of the fact that, the process is maintaining at least a minimum level of capability, its performance is highly volatile and hence should not be summarized through a single C_{pk} value.
- (2) In example 2 of Section 5, only one subgroup level observed C_{pk} value lie above the UCL, while none lie below the LCL. Since unlike Example 1, here no drastic variation is visible in the observed C_{pk} values from subgroup to subgroup, the process can be considered to be consistently performing apart from being statistically stable. Hence, its capability can very well be summarized through a single C_{pk} value.
- (3) For the simulated process, discussed in Section 4, none of the subgroup level observed C_{pk} lie beyond either of the control limits. Therefore, the process is statistically stable and consistently capable and hence its performance can be summarized through single C_{pk} value.

Through numerical examples, it has also been established that for higher sample size, say $n \geq 10$, estimation of C_{pk} , based on $\bar{X}$ –S chart information performs better than that based on $\bar{X}$ –R chart information.

Finally, apart from the point and interval estimation of C_{pk} , the corresponding testing of hypothesis is also useful from practical viewpoint. This will require derivation of the expression for the statistical distribution of the plug-in estimator of C_{pk} , based on $\bar{X}$ –R or $\bar{X}$ –S chart information and can be considered as an interesting problem to study in future.

Following are some of the salient features of the theory developed in this article:

- (1) Since the statistical distribution of the plug-in estimator of C_{pk} is complicated, very few studies are available in the literature regarding the associated bias. In the present article, a thorough study has been done in this regard and it has been observed that for sufficiently large sample size, the proposed estimators of C_{pk} has very insignificant amount of bias.
- (2) The earlier discussions on the estimation of C_{pk} using subsample information mostly considered Bayesian perspective [29]. However, often the necessary information for such study is unavailable due to various reasons like lack of infrastructure to store such data, ignorance among the competent authority and so on. The plug-in estimator of C_{pk} defined in this article does not require such information and hence is more robust from this perspective.
- (3) Till now, very few research article are available in the literature regarding process capability control chart of C_{pk} and hence the proposed charts can be used for continual assessment of the process performance.
- (4) The expressions for β -risk and ARL are discussed for practical implementation.

Disclosure statement

No potential conflict of interest was reported by the author.

References

- [1] Kotz S, Johnson N. Process capability indices – a review, 1992–2000. *J Qual Technol.* **2002**;34(1):2–19.
- [2] Pearn WL, Kotz S. Encyclopedia and handbook of process capability indices – series on quality, reliability and engineering statistics. Singapore: World Scientific; **2007**.
- [3] Lin GH, Pearn WL. Distributions of the estimated process capability index C_{pk} . *Econ Qual Control.* **2003**;18(2):263–279.
- [4] Novoa C, Leon NA. On the distribution of the usual estimator of C_{pk} and some applications in SPC. *Qual Eng.* **2009**;21(1):24–32.
- [5] Chen S-M, Hsu Y-S. Uniformly most powerful test for process capability index C_{pk} . *Qual Technol Quant Manage.* **2004**;1(2):257–269.
- [6] Lin H-C. Using normal approximation for calculating the p -value in assessing process capability index C_{pk} . *Int J Adv Manuf Technol.* **2005**;25:160–166.
- [7] Lin HC, Sheen GJ. Practical implementation of the capability index C_{pk} based on the control chart data. *Qual Eng.* **2005**;17(3):371–390.
- [8] Lin HC, Sheen GJ. An approximation approach for making decisions in assessing the capability index C_{pk} from the subsamples. *Commun Stat Simul Comput.* **2005**;34(1):191–202.
- [9] Aslam M, Azam M, Jun C-H. A mixed repetitive sampling plan based on process capability index. *Appl Math Model.* **2013**;37(24):10027–10035.
- [10] Aslam M, Wu C-W, Azam M, et al. Variable sampling inspection for resubmitted lots based on process capability index C_{pk} for normally distributed items. *Appl Math Model.* **2013**;37(3):667–675.
- [11] Aslam M, Wu C-W, Jun C-H, et al. Developing a variables repetitive group sampling plan based on process capability index C_{pk} with unknown mean and variance. *J Stat Comput Simul.* **2013**;83(8):1507–1517.
- [12] Negrin I, Parmet Y, Schechtman E. Developing a sampling plan based on C_{pk} – unknown variance. *Qual Reliab Eng Int.* **2011**;27:3–14.
- [13] Seifi S, Nezaad MSF. Variable sampling plan for resubmitted lots based on process capability index and Bayesian approach. *Int J Adv Manuf Technol.* **2016**. DOI:10.1007/s00170-016-8958-9
- [14] Wu C-W, Aslam M, Jun C-H. Variables sampling inspection scheme for resubmitted lots based on the process capability index C_{pk} . *Eur J Oper Res.* **2012**;217:560–566.

- [15] Pearn WL, Wu CH. A close form solution for the product acceptance determination based on the popular index C_{pk} . *Qual Reliab Eng Int.* **2013**;29:719–723.
- [16] Pearn WL, Liao MY. Measuring process capability based on C_{pk} with gauge measurement errors. *Microelectron Reliab.* **2005**;45:739–751.
- [17] Pearn WL, Chen KS. A Bayesian-like estimator of C_{pk} . *Comm Statist Simulation Comput.* **1996**;25(2):321–329.
- [18] Pearn WL, Chen KS. A practical implementation of the process capability index C_{pk} . *Qual Eng.* **1997**;9(4):372–383.
- [19] Pearn WL, Wu CW. Process capability assessment for index C_{pk} based on Bayesian approach. *Metrika.* **2005**;61:221–234.
- [20] Kargar M, Mashinchi M, Parchami A. A Bayesian approach to capability testing based on C_{pk} with multiple samples. *Qual Reliab Eng Int.* **2014**;30(5):615–621.
- [21] Pearn WL, Wu CC, Wu CH. Estimating process capability index C_{pk} : classical approach versus Bayesian approach. *J Stat Comput Simul.* **2015**;85:2007–2021.
- [22] Montgomery DC. *Introduction to statistical quality control*. 5th ed. New York (NY): John Wiley & Sons; **2000**.
- [23] Chatterjee M, Chakraborty AK. Distributions and process capability control charts of C_{pu} and C_{pl} using sub-group information. *Comm Statist Theory Methods.* **2015**;44:4333–4353.
- [24] Chatterjee M, Chakraborty AK. Some process capability indices for unilateral specification limits – their properties and the process capability control charts. *Comm Statist Theory Methods.* **2016**;45:7130–7160.
- [25] Spiring FA. Process capability: a total quality management tool. *Total Qual Manage.* **1995**;6(1):21–34.
- [26] Boyles RA. The Taguchi capability index. *J Qual Technol.* **1991**;23(1):17–26.
- [27] Spiring FA. A process capability/customer satisfaction approach to short-run processes. *Qual Reliab Eng Int.* **2008**;24:467–483.
- [28] Kuo H-L. Approximate tolerance limits for C_p capability chart based on range. *ProbStat Forum.* **2010**;3:145–157.
- [29] Leung BPK, Spiring F. Adjusted action limits for C_{pm} based on departures from normality. *Int J Prod Econ.* **2007**;107:237–249.
- [30] Kuo H-L, Wu C-C, Wu C-H. Approximate confidence bounds for the process capability index C_{pm} based on the range. *Inform Manage Sci.* **2006**;17(2):57–70.
- [31] Morita M, Arizono I, Nakase I, et al. Economical operation of the C_{pm} control chart for monitoring process capability index. *Int J Adv Manuf Technol.* **2009**;43:304–311.
- [32] Wu H-H. Using target costing concept in loss function and process capability indices to set up goal control limits. *Int J Adv Manuf Technol.* **2004**;24:206–213.
- [33] Wu C-W. An alternative approach to test process capability for unilateral specification with subsamples. *Int J Prod Res.* **2007**;45(22):5397–5415.
- [34] Wu C-W. Assessing process capability based on Bayesian approach with subsamples. *European J Oper Res.* **2008**;184(1):207–228.
- [35] Lord E. The use of range in place of standard deviation in the t -test. *Biometrika.* **1947**;41(1):41–67.
- [36] Leone FC, Nelson LS, Nottingham RB. The folded normal distribution. *Technometrics.* **1961**;3(4):543–550.
- [37] Woodall WH, Montgomery DC. Using ranges to estimate variability. *Qual Eng.* **2000**;13(2):211–217.
- [38] Vannman K. A unified approach to capability indices. *Statist Sin.* **1995**;5(2):805–820.
- [39] Chatterjee M, Chakraborty AK. A simple algorithm for calculating values for folded normal distribution. *J Stat Comput Simul.* **2016**;86(2):293–305.