# The Effect of Pooled and Un-pooled Variance Estimators on $C_{pm}$ When Using Subsamples

NORMA FARIS HUBELE

*Arizona State University, Tempe, AZ 85287-5906*

KERSTIN VÄNNMAN

*Luleå University of Technology, SE-971 87 Luleå, SWEDEN*

The vast majority of research on capability indices has assumed that the data consists of one large, representative sample. In practice, and in much of the quality control literature, process data are collected over time in subsamples representing rational subgroups. In this paper we examine the statistical behavior of two $C_{pm}$ estimators based on this more realistic data structure. The estimators correspond to pooled and un-pooled variance estimators. The theoretical findings are applied to hypothesis testing and power calculations. The power functions of the tests based on the two estimators are used to determine the minimum number of subsamples needed to meet a threshold requirement that power exceeds 0.80. Extensive tables of the recommended number of subsamples are provided with comments on their usage.

KEY WORDS: Capability Indices; Hypothesis Testing; Power; Sample Size.

## Introduction

A capability index is designed to quantify the relationship between the actual performance of a process and its engineering specifications. Quantifying this relationship has led to considerable debate about the adequacy of a single index to summarize process performance (see excellent summaries of this debate by Kotz and Johnson (2002) and Kotz and Lovelace (1998)). While it is necessary to recognize possible abuses of capability indices, there is no debate about their importance.

Capability analysis is an important step in implementing a control system. Whether one chooses to compute an index or make some other comparison between the process output and the engineering specifications, the analysis should be performed using data from a stable process. Typically, process stability is assessed by collecting subsamples at regular intervals and plotting subsample statistics on control charts. Once the charts show a reasonable degree of stability, process capability can be assessed.

We consider two capability indices frequently used in industry today,

$$C_p = \frac{USL - LSL}{6\sigma} \quad \text{and}$$

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

where $[LSL, USL]$ is the specification interval, $\mu$ is the process mean, and $\sigma$ is the process standard deviation. Since statistical control has been established using subsamples (i.e., small samples or rational subgroups), it is tempting to use estimators of the unknown process parameters $\mu$ and $\sigma$ based on

Dr. Hubele is a Professor in the Department of Industrial Engineering. Her email address is hubele@asu.edu.

Dr. Vännman is a Professor in the Department of Mathematics. Her email address is kerstin.vannman@sm.luth.se.

these subsamples. Most results in the literature assume these parameters are estimated from one large representative sample of process data. Exceptions include the work by Kirmani, Kocherlakota, and Kocherlakota (1991), in which the distribution of an estimator of $C_p$ based on the sample standard deviation of the subsamples was studied, and by Li, Owen, and Borrego (1990), in which the distributions of estimators of $C_p$ and $C_{pk}$ based on the range of subsamples were studied.

Vännman & Hubele (2003) studied the distribution of estimators of a general class of capability indices, including $C_p$, $C_{pk}$, and $C_{pm}$, when the process mean is estimated using the grand average and the process variance is estimated using the pooled variance. The method assumes that the estimation occurs after numerous subsamples have been collected, plotted on a control chart, and the process has been deemed to be in control.

In this paper, we focus attention on the capability index $C_{pm}$, which is gaining attention due to its sensitivity to a process target (again, see Kotz and Johnson (2002)). We let $d = (USL - LSL)/2$, i.e., half the length of the specification interval, and then define $C_{pm}$ as

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{d}{3\sqrt{\sigma^2 + (\mu - T)^2}}. \quad (2)$$

Here, we present different useful forms of the distribution of an estimator of $C_{pm}$, when the estimates of the process parameters $\mu$ and $\sigma$ are computed using subsamples from an underlying normal distribution. We assume that the process has been deemed reasonably stable using an $\overline{X}$-chart and an $S$-chart and that the underlying data are independently, identically distributed. We consider an un-pooled and pooled estimator of the variance from subsamples, and give the sampling distributions of the corresponding estimators of $C_{pm}$. The un-pooled variance estimator is equivalent to the traditional "overall" or "long-term" variance estimator, whereas the pooled estimator is based on the control-chart related "within" and "short-term" variance estimator. Under the assumption that the data are independently, identically, and normally distributed with constant mean and variance, we consider this distinction in terminology as artificial when the process is under statistical control (Rodriguez (2002)).

In the third section of this paper, the sampling distributions of $C_{pm}$ estimators, derived from the pooled and un-pooled variance estimators, are used to demonstrate a procedure for hypothesis testing and power computations. As can be seen, to attain a specified level of power the un-pooled estimator requires fewer subsamples than the pooled estimator. This result simply reflects the greater efficiency of the un-pooled over the pooled estimator. We use plots of the probability density functions to provide insight into the effect of process changes on the power of testing procedures based on the two $C_{pm}$ estimators. We present extensive tables of the recommended number of subsamples for given subsample sizes. The tables also offer the practitioner the opportunity to contrast the sampling requirements of the two different estimators to obtain a fixed power. We conclude the paper with a brief discussion of recommendations for using $C_{pm}$ with subsamples.

## Estimators of $C_{pm}$ Using Subsamples and Their Distributions

We assume that the characteristic of the process is normally distributed and that we have $m$ subsamples, each of the same sample size $n$. For each $i$, $i = 1$, 2, ..., $m$, we let $X_{ij}$, $j = 1, 2, ..., n$ be a random sample from a normal distribution with mean $\mu$ and variance $\sigma^2$ of the quality characteristic of interest. We consider the case when using an $\overline{X}$-chart together with an $S$-chart in quality control. Using these charts, we assume that the process is in a state of statistical process control during the sampling time period. Then, for each subsample $i$, the sample mean and sample variance are defined as

$$\overline{X}_i = \frac{\sum_{j=1}^{n} X_{ij}}{n} \quad \text{and} \quad S_i^2 = \frac{\sum_{j=1}^{n}\left(X_{ij} - \overline{X}_i\right)^2}{n-1}. \quad (3)$$

As an estimator of $\mu$ we use the overall sample mean, which is an unbiased estimator, i.e.,

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{m} \overline{X}_i}{m}. \quad (4)$$

We now consider two ways to compute the variance estimator used in estimating $C_{pm}$ in Equation (2).

**Pooled Variance Estimator**

One estimator of $\sigma^2$ is the pooled variance estimator, defined as

$$s_p^2 = \frac{1}{mn}\sum_{i=1}^{m}(n-1)S_i^2. \quad (5)$$

In this case, the estimator of $C_{pm}$ will be denoted by

$$\widehat{C}_{pm,p} = \frac{d}{3\sqrt{s_p^2 + \left(\overline{\overline{X}} - T\right)^2}}. \quad (6)$$

Following the derivations used by others (Boyles (1991), Chan, Cheng, and Spiring (1988), and Vännman (1997)) for obtaining the sampling distribution of $C_{pm}$ when using only a single sample, we now state the sampling distribution when using $m$ subsamples of constant size $n$. Let $F_\zeta$ denote the cumulative distribution function of $\zeta$, where $\zeta$ is distributed according to a non-central $\chi^2$-distribution with $m(n-1)+1$ degrees of freedom and non-centrality parameter $\lambda$, where

$$\lambda = \frac{mn(\mu - T)^2}{\sigma^2}. \quad (7)$$

Then, the cumulative distribution function of $\widehat{C}_{pm,p}$ can be expressed as

$$F_p(x; m, n, \mu, \sigma) = 1 - F_\zeta\left(\frac{d^2mn}{9x^2\sigma^2}\right), \quad x > 0 \quad (8)$$

(see Vännman and Hubele (2003) for derivation). The result in Equation (8) can alternatively be expressed as

$$\left(\widehat{C}_{pm,p}\right)^2 \sim \frac{d^2mn}{9\sigma^2\zeta}.$$

This in turn can be written in the following useful form,

$$\left(\frac{C_{pm}}{\widehat{C}_{pm,p}}\right)^2 \sim \frac{\zeta}{mn+\lambda}, \quad (9)$$

where $\lambda$ is given in Equation (7).

To calculate values of the cumulative distribution function of $\widehat{C}_{pm,p}$, software with numerical integration of the non-central $\chi^2$-distribution is needed. If such software is not available, we can derive an approximate expression for the cumulative distribution function given in Equation (8) based on a central $\chi^2$-distribution using the simple Patnaik's approximation (see, e.g., Johnson, Kotz, and Balakrishnan (1995, p. 462)). Such an approximation was suggested by Boyles (1991) for the single sample estimator. The non-central $\chi^2$-distributed random variable $\zeta$ can be approximated by a central $\chi^2$-distributed random variable, $\chi^2_{f_p}$, with $f_p$ degrees of freedom, according to

$$\zeta \dot{\sim} g_p\chi^2_{f_p},$$

where

$$f_p = \frac{(mn + \lambda^2 - (m-1))^2}{mn + 2\lambda - (m-1)}, \quad \text{and} \quad (10)$$

$$g_p = \frac{mn + 2\lambda - (m-1)}{mn + \lambda - (m-1)}, \quad (11)$$

and $\lambda$ is given in Equation (7).

Based on this approximation we have

$$F_p(x; m, n, \mu, \sigma) \approx 1 - F_{\chi^2_{f_p}}\left(\frac{d^2mn}{g_p 9x^2\sigma^2}\right), \quad x > 0, \quad (12)$$

where $F_{\chi^2_{f_p}}$ represents the cumulative distribution function of $\chi^2_{f_p}$. As in Equation (9), this result can alternatively be expressed as

$$\left(\frac{C_{pm}}{\widehat{C}_{pm,p}}\right)^2 \dot{\sim} \frac{g_p\chi^2_{f_p}}{mn+\lambda}. \quad (13)$$

**Un-pooled Variance Estimator**

While it is customary to use the pooled variance estimator when we have independent subsamples, it is also possible to compute an un-pooled estimator of $\sigma^2$, since we can obtain $\sum_{j=1}^{n} X_{ij}^2$ from Equation (3) through the relation

$$\sum_{j=1}^{n}\left(X_{ij} - \overline{X}_i\right)^2 = \sum_{j=1}^{n} X_{ij}^2 - n\overline{X}_i^2$$

without requiring each individual observation $X_{ij}$. Hence, given the summary statistics $\overline{X}_i$ and $S_i^2$, we can form the following un-pooled estimator of $\sigma^2$

$$s_u^2 = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij} - \overline{\overline{X}}\right)^2 = \frac{1}{mn}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}X_{ij}^2 - mn\,\overline{\overline{X}}^2\right). \tag{14}$$

Here, the corresponding estimator of $C_{pm}$ is denoted by

$$\widehat{C}_{pm,u} = \frac{d}{3\sqrt{s_u^2 + \left(\overline{\overline{X}} - T\right)^2}}. \tag{15}$$

To obtain the distribution of the estimator given in Equation (15), we let $F_\xi$ be the cumulative distribution function of $\xi$, where $\xi$ is distributed according to a non-central $\chi^2$-distribution with $mn$ degrees of freedom and non-centrality parameter $\lambda$ remains as defined above in Equation (7). Then, the cumulative distribution of $\widehat{C}_{pm,u}$ can be expressed as

$$F_u(x; m, n, \mu, \sigma) = 1 - F_\xi\left(\frac{d^2 mn}{9x^2\sigma^2}\right), \quad x > 0. \tag{16}$$

As with Equation (9), this result can be written as

$$\left(\frac{C_{pm}}{\widehat{C}_{pm,u}}\right)^2 \sim \frac{\xi}{mn + \lambda}.$$

As for the pooled case, the non-central $\chi^2$-distributed random variable $\xi$ can be approximated by a central $\chi^2$-distributed random variable $\chi^2_{f_u}$, with $f_u$ degrees of freedom, according to the following:

$$\xi \mathrel{\dot\sim} g_u \chi^2_{f_u},$$

where

$$f_u = \frac{(mn + \lambda^2)^2}{mn + 2\lambda}, \quad \text{and} \tag{17}$$

$$g_u = \frac{mn + 2\lambda}{mn + \lambda}, \tag{18}$$

and where $\lambda$ is given in Equation (7).

Based on this approximation, we have

$$F_u(x; m, n, \mu, \sigma) \approx 1 - F_{\chi^2_{f_u}}\left(\frac{d^2 mn}{g_u 9x^2\sigma^2}\right), \quad x > 0, \tag{19}$$

where $F_{\chi^2_{f_u}}$ denotes the cumulative distribution function of $\chi^2_{f_u}$. As with Equation (13), this result can be expressed as

$$\left(\frac{C_{pm}}{C_{pm,u}}\right)^2 \mathrel{\dot\sim} \frac{g_u \chi^2_{f_u}}{mn + \lambda}. \tag{20}$$

We note that the cumulative distribution function in Equation (16) is identical to the usual cumulative distribution function given elsewhere, when a single sample of size $mn$ is used to estimate $C_{pm}$ (see Kotz and Johnson (1993)). The probability density functions of the estimators of $C_{pm}$ are easily obtained by taking the derivative with respect to $x$ of either $F_p(x; m, n, \mu, \sigma)$ or $F_u(x; m, n, \mu, \sigma)$.

In this paper, we address the situation where an $S$-chart has been used in conjunction with an $\overline{X}$-chart to monitor the process and remove assignable causes. Under these conditions, we assume that we have a stable underlying normal distribution with constant mean and variance. Here, the use of a pooled or unpooled estimator of the process variance is quite sensible. In practice, the modeling assumption of independently, identically distributed normal random observations with constant variance may be suspect. In the case of non-constant variance, one might examine the possibility of a more complex variance structure including separate components representing the within-subsample and between-subsample variability, as determined using ANOVA mean squares. In the case where the assumption of independent, identically distributed random data is weakly violated, it may still be useful to make this assumption for sensitivity analysis and power investigations, as will be done in this paper. Consequently, for the purposes of this paper, we assume that Equations (5) and (14) are estimators for the same unknown parameter $\sigma^2$ from a stable process with independently, identically distributed normal random observations.

## Hypothesis Testing

When using a process capability index, a process is defined as capable if the index exceeds a certain value $k > 0$. Some commonly used values are $k = 4/3$, 1.5, and 1.6. In order to determine if a process is capable or not, we need a decision rule based on an estimator. To obtain such a decision rule, we consider a hypothesis test with the null hypothesis $H_0: C_{pm} \leq k$ and the alternative hypothesis given by $H_1: C_{pm} > k$. Such an approach was also used by

Cheng (1992) in his investigations of the single sample case, with $\mu = T$. In contrast to Cheng's estimator based on a single sample, we consider the two test statistics based on the two estimators defined in Equations (6) and (15). Furthermore, we do not restrict our investigation to the case where $\mu = T$.

For the moment, we use the notation $\widehat{C}_{pm}$ to denote the estimator, and only add the additional subscripts $p$ and $u$ when needed to distinguish between the two estimators. The null hypothesis is rejected whenever $\widehat{C}_{pm} > c_\alpha$, where the constant $c_\alpha$ is determined so that the significance level of the test is $\alpha$ for the corresponding estimator. For given values of $\alpha$, $m$, and $n$, the decision rule is that the process is considered capable if $\widehat{C}_{pm} > c_\alpha$ and non-capable if $\widehat{C}_{pm} \le c_\alpha$.

In order to find the critical value $c_\alpha$, we need to calculate $P(\widehat{C}_{pm} > c_\alpha)$, when $C_{pm} = k$, using the results of the previous section. We see from Equations (8), (12), (16), and (19) that the distribution of $\widehat{C}_{pm}$ depends on $\mu$ and $\sigma$ but not solely through $C_{pm}$. Hence, for a given value of $C_{pm}$ there is not a unique probability $P(\widehat{C}_{pm} > c_\alpha)$. To calculate this probability, we need to have values for both $\mu$ and $\sigma$. We consider the notation

$$\delta = \frac{\mu - T}{d} \quad \text{and} \quad \gamma = \frac{\sigma}{d}, \quad \text{where } |\delta| \le 1. \qquad (21)$$

When $C_{pm} = k$, we have the following relation between $\gamma$ and $\delta$:

$$\gamma = \sqrt{\frac{1}{9k^2} - \delta^2}, \quad |\delta| \le \frac{1}{3k}. \qquad (22)$$

When we plot $\gamma$ as a function of $\delta$, we obtain a semi-circle. See Figure 1 for $k = 4/3$ and $k = 5/3$. Values of the process parameters $\mu$ and $\sigma$ which give $(\delta, \gamma)$-values inside the region bounded by the semi-circle $C_{pm} = k$ and the $\delta$-axis will give rise to a $C_{pm}$-value larger than $k$ (i.e., a capable process). We can call this region of the $(\delta, \gamma)$-plot the *capability region*. Furthermore, values of $\mu$ and $\sigma$ which give $(\delta, \gamma)$-values outside this region will give a $C_{pm}$-value smaller than $k$ (i.e., a non-capable process). For more discussion regarding capability regions for a general class of capability indices, see Deleryd and Vännman (1999).

When $(\mu, \sigma)$, i.e., $(\delta, \gamma)$, moves along the semi-circle, the distribution of $\widehat{C}_{pm}$ will vary, so the probability $P(\widehat{C}_{pm} > c_\alpha)$, given that $C_{pm} = k$, will vary. For some examples, see Appendix A and Vännman and Hubele (2001, 2003).

Since our null hypothesis $H_0 : C_{pm} \le k$ is composite, we need to investigate $P(\widehat{C}_{pm} > c_\alpha)$ for all values of $(\mu, \sigma)$ for which $C_{pm} \le k$ in order to obtain the critical value $c_\alpha$. This null hypothesis can be reduced to $H_0 : C_{pm} = k$ (see Vännman and Hubele (2001, 2003)). However, this is still a composite hypothesis. Hence, we need to calculate $P(\widehat{C}_{pm} > c_\alpha)$, given $C_{pm} = k$, along the semi-circle in Equation (22) to find the value of $\delta$ that maximizes this probability. Using this value of $\delta$ we can then derive our critical value for the two estimators. The variation of the probability $P(\widehat{C}_{pm} > c_\alpha)$ for a given value of $C_{pm} = k$ along the semi-circle leads to some special reasoning when deriving the critical value as well as when studying the power.

### Pooled Variance Estimator

If we combine the results in Equation (8), (21), and (22) for the estimator $\widehat{C}_{pm}$ in Equation (6), we find that, given $C_{pm} = k$, we can express $P(\widehat{C}_{pm,p} > x)$ as

$$P\left(\widehat{C}_{pm,p} > x \mid C_{pm} = k\right) = F_\zeta\left(\frac{k^2 mn}{x^2(1 - 9k^2\delta^2)}\right), \quad x > 0, \qquad (23)$$

where $\zeta$ is distributed according to a non-central $\chi^2$-distribution with $m(n-1)+1$ degrees of freedom and non-centrality parameter $\lambda$, where

$$\lambda = \frac{9k^2\delta^2 mn}{1 - 9k^2\delta^2}. \qquad (24)$$

As in the previous section, we can apply Patnaik's approximation to obtain an approximate expression for the results in Equation (23) based on a central $\chi^2$-distribution. Hence, we get

$$\zeta \dot{\sim} g_p \chi^2_{f_p},$$

where the expression for $\lambda$ in Equation (24) has to be included in the expressions for the degrees of freedom $f_p$ and the constant $g_p$ in Equations (10) and (11), respectively.

Combining the results in Equations (12), (21), and (22), we find that, given $C_{pm} = k$, we can express $P\left(\widehat{C}_{pm,p} > x\right)$ approximately as

$$P\left(\widehat{C}_{pm,p} > x \mid C_{pm} = k\right) \approx F_{\chi^2_{f_p}}\left(\frac{k^2 mn}{g_p x^2(1 - 9k^2\delta^2)}\right), \quad x > 0. \qquad (25)$$

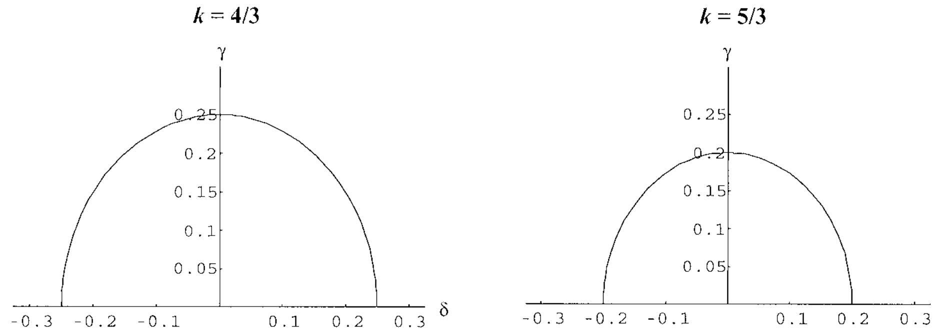


FIGURE 1. The Semi-Circle Obtained from Equation (22) for $k = 4/3$ and $k = 5/3$. The Region Bounded by the Semi-Circle and the $\delta$-axis is the Capability Region.

We see that the probability given in Equation (25) is a function of $\delta$, since $f_p$, $g_p$, and the argument of the distribution function depend upon $\delta$. In fact, the probability is symmetric around $\delta = 0$. Hence, it is sufficient to explore this probability only for values $\delta \geq 0$ along the semi-circular capability level curve of $C_{pm}$, as discussed and illustrated by Vännman and Hubele (2001, 2003). Arizono et al. (1997) showed that this probability, as a function of $\delta$, has its maximum at $\delta = 0$. From Equation (24) we see that when $\delta = 0$, the non-centrality parameter $\lambda = 0$. Hence, the critical value can be determined using a central $\chi^2$-distribution. From Equation (23) we obtain

$$c_{\alpha,p} = k_0 \sqrt{\frac{mn}{\chi^2_{\alpha,m(n-1)+1}}}, \tag{26}$$

where $\chi^2_{\alpha,m(n-1)+1}$ is the $\alpha$ quantile from the $\chi^2$-distribution with $m(n-1)+1$ degrees of freedom. We note that when $\delta = 0$, the Patnaik's approximation in Equation (25) equals the expression in Equation (23), and thus generates the same critical value.

The probability of a Type I error, i.e., the probability of saying that a process is capable when it is not, is controlled by setting the significance level of the test, $\alpha$. The probability of a Type II error, $\beta$, i.e., the probability of saying a process is not capable when it is capable, can be explored by considering true values of $C_{pm}$ greater than $k_0$. Of special interest is the power of the test, $1 - \beta$, i.e., the probability that the process is considered capable when $C_{pm} = k_1 > k_0$. For the power, when using the pooled variance estimator, we use the notation

$$\begin{aligned} \text{Pow}_p(\delta; m, n, \alpha, k_0, k_1) & \\ &= P\left(\widehat{C}_{pm,p} > c_{\alpha,p} \mid C_{pm} = k_1\right) \\ &= F_{\zeta}\left(\frac{k_1^2 \chi^2_{\alpha,m(n-1)+1}}{k_0^2\left(1 - 9k_1^2\delta^2\right)}\right) \\ &\approx F_{\chi^2_{f_p}}\left(\frac{k_1^2 \chi^2_{\alpha,m(n-1)+1}}{g_p k_0^2\left(1 - 9k_1^2\delta^2\right)}\right), \quad x > 0. \end{aligned} \tag{27}$$

This probability varies as $(\mu, \sigma)$, i.e., $(\delta, \gamma)$, moves along the semi-circle defined by $C_{pm} = k_1$. Hence, for a given value of $C_{pm} = k_1$, the power depends on $\delta$. We also note that the argument and the degrees of freedom depend on $m$ and $n$ in Equation (27) through the functional relationship in Equations (10) and (11).

Consider the example when $k_0 = 4/3$ and $k_1 = 1.9$. We illustrate the behavior of the power when $n = 4$ for the situations: (a) when $m = 10$ and $\alpha = 0.10$; and (b) when $m = 14$ and $\alpha = 0.05$. These values of $m$ are the minimum numbers needed to have a power of at least 0.80 for all values of $\delta$ along the semi-circle defined by $C_{pm} = 1.9$. The corresponding critical values, obtained from Equation (26), for these two situations are $c_{0.10,p} = 1.8215$ and $c_{0.05,p} = 1.8540$, respectively. In Figures 2a and 2b, the corresponding power graph, when $C_{pm} = k_1 = 1.9$, is plotted as a function of $\delta$. The graphs demonstrate that the power decreases as $\delta$ moves away from 0 and then increases. In particular, the minimum power occurs when $\delta$ is close to 0.17, i.e., when $\mu$ is away from the target value and, at the same time, $\sigma$ is very small. An explanation of the effect of $\delta$ on the behavior of the power curve is given in Appendix A.

Vännman and Hubele (2001, 2003) considered numerous additional illustrations and demonstrated that, for a fixed value of $n$ and significance level $\alpha$, it

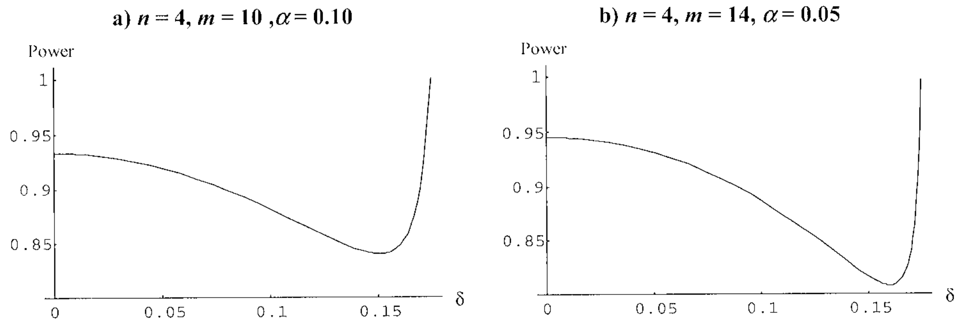


FIGURE 2. The Power for $C_{pm} = k_1 = 1.9$, as a Function of $\delta = (\mu - T)/d$, When Using the Pooled Estimator and $k_0 = 4/3$.

is possible to bound power from below by increasing $m$. We use this fact to provide recommendations for determining the minimum number of subsamples for the pooled case in the section on sampling requirements below.

### Un-pooled Variance Estimator

Similarly, using the estimator $\widehat{C}_{pm,u}$ in Equation (15), together with the results in Equations (16), (21), and (22) we find that

$$P\left(\widehat{C}_{pm,u} > x \mid C_{pm} = k\right) = F_\xi\left(\frac{k^2mn}{x^2(1-9k^2\delta^2)}\right), \quad x > 0, \tag{28}$$

where $\xi$ is distributed according to a non-central $\chi^2$-distribution with $mn$ degrees of freedom and the non-centrality parameter $\lambda$ remains as defined in Equation (24). The corresponding Patnaik's approximation is obtained through the relation

$$\xi \,\dot{\sim}\, g_u \chi^2_{f_u},$$

where the expression for $\lambda$ in Equation (24) has to be included in the expressions for $f_u$ and $g_u$ in Equations (17) and (18), respectively. Hence, we have

$$P\left(\widehat{C}_{pm,u} > x \mid C_{pm} = k\right) \approx F_{\chi^2_{f_u}}\left(\frac{k^2mn}{g_u x^2(1-9k^2\delta^2)}\right), \quad x > 0. \tag{29}$$

The critical value associated with the estimator $\widehat{C}_{pm,u}$ and the null hypothesis $H_0 : C_{pm} = k_0$ can be determined using a central $\chi^2$-distribution, as in the pooled case. From Equation (28) we obtain

$$c_{\alpha,u} = k_0\sqrt{\frac{mn}{\chi^2_{\alpha,mn}}}, \tag{30}$$

where the degrees of freedom are now $mn$. For the power, when $C_{pm} = k_1 > k_0$, we use the notation

$$\begin{aligned} \mathrm{Pow}_u(\delta; m, n, \alpha, k_0, k_1) &= P\left(\widehat{C}_{pm,u} > c_{\alpha,u} \mid C_{pm} = k_1\right) \\ &= F_\xi\left(\frac{k_1^2\chi^2_{\alpha,mn}}{k_0^2(1-9k_1^2\delta^2)}\right) \\ &\approx F_{\chi^2_{f_u}}\left(\frac{k_1^2\chi^2_{\alpha,mn}}{g_u k_0^2(1-9k_1^2\delta^2)}\right), \quad x > 0. \end{aligned} \tag{31}$$

This probability, once again, varies as $(\mu, \sigma)$, i.e., $(\delta, \gamma)$, moves along the semi-circle defined by $C_{pm} = k_1$. Also, we observe that both the argument and the degrees of freedom depend upon $m$ and $n$ in Equation (31) through the functional relationship of Equations (17) and (18). Hence, the effect of $m$ and $n$ on the power calculation in Equation (31) is different than the effect on the power calculation in the pooled case in Equation (27).

Consider, as in the pooled case, the example when $k_0 = 4/3$ and $k_1 = 1.9$. Here, we illustrate the behavior of the power when $n = 4$ for two different situations: (a) when $m = 5$ and $\alpha = 0.10$; and (b) when $m = 7$ and $\alpha = 0.05$. These values of $m$ are the minimum numbers needed to have a power of at least 0.80 for all values of $\delta$ along the semi-circle defined by $C_{pm} = 1.9$. Consequently, the values in this illustration are purposely chosen to be different from the pooled case since larger values of $m$ do not yield good illustrative examples. The corresponding critical values for these two situations are $c_{0.10,u} = 1.6904$ and $c_{0.05,u} = 1.7148$, respectively. In Figures 3a and 3b, the corresponding power graphs, as functions of $\delta$, are plotted. Unlike the pooled case, both of these graphs demonstrate that the power monotonically increases as $\delta$ increases. Furthermore, the minimum

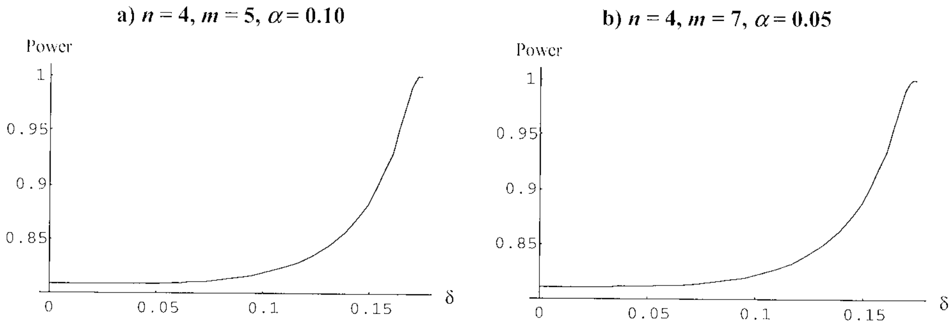


FIGURE 3. The Power for $C_{pm} = k_1 = 1.9$, as a Function of $\delta = (\mu - T)/d$, When Using the Un-Pooled Estimator and $k_0 = 4/3$.

power always occurs at $\delta = 0$, i.e., at $\mu = T$. Such is the behavior of all other situations investigated for the un-pooled case. Again, an explanation of the behavior of the power curve is given in Appendix A.

We use the reasoning above to provide recommendations for determining the minimum number of subsamples for the un-pooled case in the section on sampling requirements below.

## Comparing the Pooled and Un-pooled Cases

For convenience, we combine the power graphed in Figures 2 and 3 into Figure 4. The illustrations for these two cases are chosen such that the number of subsamples for a given subsample size meet the minimum power requirement of 0.80. We see clearly in Figure 4 that although the minimum power is approximately the same, the curve shape is quite different, depending on which estimator is used. For the un-pooled case, the power always achieves its minimum at $\delta = 0$. In the pooled case, however, the minimum is achieved at different $\delta$ values in the interval $[0, k_1/3]$, depending on the values of $m$, $n$, $\alpha$, $k_0$, and $k_1$. The minimum often occurs when $\delta$ is away from 0, i.e., $\mu$ is away from the target and, at the same time, $\sigma$ is small. These kinds of graphs are used to identify adequate sampling requirements for stated significance and minimum power as described below. See Vännman and Hubele (2003) for further comparisons of the power in the case when $m$ and $n$ are kept equal in the pooled and un-pooled case.

## Sampling Requirements

In practice, the subsample size $n$, used to monitor and control a process, is determined by the concept of rational subgrouping. In this regard, consider $n$ to be given. For fixed $n$, we have found that the minimum power, when using either estimator, increases as $m$ increases. In the following discussion, we set a minimum power requirement of 0.80.

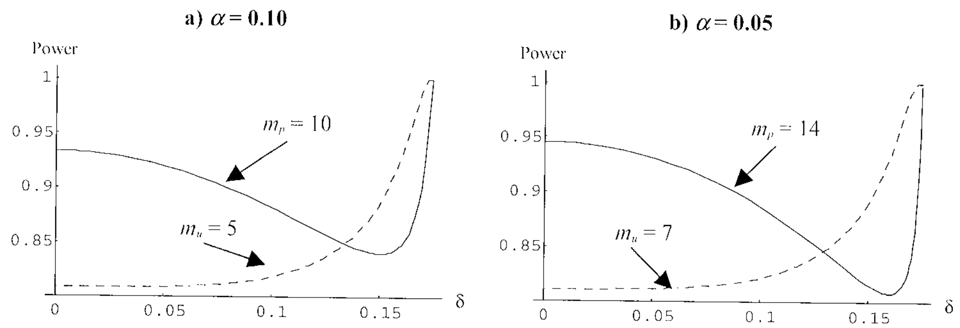


FIGURE 4. The Power for $C_{pm} = k_1 = 1.9$, as a Function of $\delta = (\mu - T)/d$ for the Pooled and the Un-Pooled Cases, When $k_0 = 4/3$, $n = 4$. a) The Continuous Line Corresponds to the Pooled Case with $m = 10$. The Dashed Line Corresponds to the Un-Pooled Case with $m = 5$. b) The Continuous Line Corresponds to the Pooled Case with $m = 14$. The Dashed Line Corresponds to the Un-Pooled Case with $m = 7$.

Now the question to be addressed is, what is the minimum number of subsamples needed to meet the minimum power requirement? In other words, what value of $m$ should be used to test the capability of a process, given subsample size $n$, capability index threshold $k_0$, and significance level $\alpha$?

### Un-pooled and Pooled Cases

For a specified $k_1$, we consider the following notation. We let $m_u$ represent the minimum number of subsamples with the un-pooled estimator and $m_p$ represent the minimum with the pooled estimator. As seen in the following tables, $m_u \leq m_p$. Furthermore, although we place a threshold requirement on power, the power of the test for a given $\delta$, using either estimator, may exceed 0.80 depending on the actual value of $\delta$, as was seen in Figures 2, 3, and 4.

As is shown in Appendix B, there is a closed-form expression for $m_u$. The expression is based on the Wilson-Hilferty approximation of the $\chi^2$-distribution using the standard normal. We have $m_u$, the smallest integer for which

$$m_u \geq \frac{1}{n}\left(A + \sqrt{A^2 + \frac{2}{9}}\right)^2, \tag{32}$$

where

$$A = \frac{k_0^{2/3} z_{1-\beta} - k_1^{2/3} z_\alpha}{k_1^{2/3} - k_0^{2/3}} \cdot \frac{1}{3\sqrt{2}}$$

and the quantities $z_\alpha$ and $z_{1-\beta}$ are the $\alpha$ and $(1-\beta)$ quantiles of the N(0,1) distribution, respectively.

In Tables 1 through 9, we present $m_u$ and $m_p$ for $k_0 = 4/3$, 1.5, and 1.6. For each of these null hypothesis values, we consider $n = 4$, 5, 6, 7, 8, 9, and 10 and significance levels $\alpha = 0.10$, 0.05, and 0.01. Depending on the value of $k_0$, four $k_1$ values greater than $k_0$ are considered.

To illustrate the use of these tables, suppose we have $n = 4$, $k_0 = 4/3$, $k_1 = 1.9$, and $\alpha = 0.10$. From Table 1, with the un-pooled estimator $m_u = 5$, whereas with the pooled estimator $m_p = 10$. For insight into the power curves of these two cases, the reader is referred to Figures 2a, 3a, and 4a. Furthermore, if we decrease the significance level to $\alpha = 0.05$, then from Table 2 we find that $m_u = 7$ and $m_p = 14$. As can be seen in Table 3, for $\alpha = 0.01$, $m_u = 12$ and $m_p = 26$. The pattern of increasing $m_u$ and $m_p$ is consistent across all cases as $\alpha$ decreases.

When $k_0$ and $\alpha$ are fixed, and we examine the $m_u$ and $m_p$ values across a row for a given $k_1$, we see two patterns. One pattern is the decreasing number of subsamples required for each of the estimators. The second is that the difference between $m_u$ and $m_p$ decreases as $n$ increases. When $n$ is as small as 4 or 5 we can see rather large differences; however, when $n$ is 9 or 10 the difference is very small or non-existent. These patterns support the intuitive notions that fewer subsamples are needed as the subsample size increases and that the un-pooled estimator is more efficient than the pooled estimator, especially when $n$ is small.

As the difference between $k_0$ and $k_1$ increases, the required number of subsamples decreases. In a few instances, when the significance level is 0.01 and $k_0$ and $k_1$ are very close together, as in Tables 3, 6, and 9, $m_u$ and $m_p$ will exceed the practical limits of 100 subsamples. In those instances, if it is necessary to detect such a small difference, then we recommend a larger significance level that requires fewer subsamples, as given in the other tables for $\alpha = 0.10$ and 0.05.

## Concluding Remarks

Capability indices are designed to summarize process performance with respect to engineering requirements. Process performance is quantified by collecting data on key quality characteristics and forming summary statistics. The vast majority of the theory and methods supporting the use of capability indices has centered on the assumption that the process data are collected using one large representative sample (see, e.g., Kotz and Johnson (2002)). In practice and in much of the quality control literature, however, the process performance is monitored and controlled by periodically collecting subsamples of data, e.g., based on the concepts of rational subgrouping. In this paper we depart from the typical assumption of estimators of $C_{pm}$ and examine the statistical behavior of the estimator of $C_{pm}$ when it is based on numerous subsamples of data.

We have provided the distribution functions of two different $C_{pm}$ estimators, based on pooled and un-pooled variance estimators. These theoretical findings are then used to illustrate the application to hypothesis testing and computations of the power of the test. Finally, the power of these two estimators is used to generate the minimum sampling requirements for various subsample sizes.

TABLE 1. The Minimum Number of Subsamples, $m$, for Power $\geq 0.80$, When $k_0 = 4/3$ and $\alpha = 0.10$

| | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 1.7$ | | | | | | | |
| $m_u$ | 11 | 9 | 7 | 6 | 6 | 5 | 5 |
| $m_p$ | 35 | 17 | 11 | 9 | 7 | 6 | 5 |
| $k_1 = 1.8$ | | | | | | | |
| $m_u$ | 7 | 6 | 5 | 4 | 4 | 3 | 3 |
| $m_p$ | 16 | 9 | 7 | 5 | 5 | 4 | 3 |
| $k_1 = 1.9$ | | | | | | | |
| $m_u$ | 5 | 4 | 4 | 3 | 3 | 3 | 2 |
| $m_p$ | 10 | 6 | 5 | 4 | 3 | 3 | 3 |
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 4 | 4 | 3 | 3 | 2 | 2 | 2 |
| $m_p$ | 7 | 5 | 4 | 3 | 3 | 2 | 2 |

TABLE 2. The Minimum Number of Subsamples, $m$, for Power $\geq 0.80$, When $k_0 = 4/3$ and $\alpha = 0.05$

| | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 1.7$ | | | | | | | |
| $m_u$ | 14 | 12 | 10 | 8 | 7 | 7 | 6 |
| $m_p$ | 55 | 26 | 17 | 12 | 10 | 8 | 7 |
| $k_1 = 1.8$ | | | | | | | |
| $m_u$ | 10 | 8 | 7 | 6 | 5 | 5 | 4 |
| $m_p$ | 24 | 14 | 10 | 8 | 6 | 5 | 5 |
| $k_1 = 1.9$ | | | | | | | |
| $m_u$ | 7 | 6 | 5 | 4 | 4 | 4 | 3 |
| $m_p$ | 14 | 9 | 7 | 5 | 4 | 4 | 4 |
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 6 | 5 | 4 | 4 | 3 | 3 | 3 |
| $m_p$ | 10 | 7 | 5 | 4 | 3 | 3 | 3 |

TABLE 3. The Minimum Number of Subsamples, $m$, Needed for Power $\geq 0.80$, When $k_0 = 4/3$ and $\alpha = 0.01$

| | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 1.7$ | | | | | | | |
| $m_u$ | 24 | 19 | 16 | 14 | 12 | 11 | 10 |
| $m_p$ | >100 | 47 | 30 | 22 | 17 | 14 | 12 |
| $k_1 = 1.8$ | | | | | | | |
| $m_u$ | 16 | 13 | 11 | 9 | 8 | 7 | 7 |
| $m_p$ | 44 | 25 | 17 | 13 | 11 | 9 | 8 |
| $k_1 = 1.9$ | | | | | | | |
| $m_u$ | 12 | 10 | 8 | 7 | 6 | 6 | 5 |
| $m_p$ | 26 | 16 | 11 | 9 | 8 | 7 | 6 |
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 9 | 8 | 6 | 6 | 5 | 4 | 4 |
| $m_p$ | 17 | 11 | 8 | 7 | 6 | 5 | 4 |

TABLE 4. The Minimum Number of Subsamples, $m$, Needed for Power $\geq 0.80$, When $k_0 = 1.5$ and $\alpha = 0.10$

| | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 1.9$ | | | | | | | |
| $m_u$ | 11 | 9 | 8 | 7 | 6 | 5 | 5 |
| $m_p$ | 40 | 19 | 12 | 9 | 8 | 6 | 6 |
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 8 | 6 | 5 | 5 | 4 | 4 | 3 |
| $m_p$ | 18 | 10 | 7 | 6 | 5 | 4 | 4 |
| $k_1 = 2.1$ | | | | | | | |
| $m_u$ | 6 | 5 | 4 | 4 | 3 | 3 | 3 |
| $m_p$ | 11 | 7 | 5 | 4 | 4 | 3 | 3 |
| $k_1 = 2.2$ | | | | | | | |
| $m_u$ | 5 | 4 | 3 | 3 | 3 | 2 | 2 |
| $m_p$ | 8 | 5 | 4 | 3 | 3 | 3 | 2 |

TABLE 5. The Minimum Number of Subsamples, $m$, Needed for Power $\geq 0.80$, When $k_0 = 1.5$ and $\alpha = 0.05$

| | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 1.9$ | | | | | | | |
| $m_u$ | 15 | 12 | 10 | 9 | 8 | 7 | 6 |
| $m_p$ | 51 | 27 | 18 | 13 | 11 | 9 | 8 |
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 11 | 9 | 7 | 6 | 6 | 5 | 5 |
| $m_p$ | 28 | 15 | 11 | 8 | 7 | 6 | 5 |
| $k_1 = 2.1$ | | | | | | | |
| $m_u$ | 8 | 7 | 6 | 5 | 4 | 4 | 4 |
| $m_p$ | 17 | 10 | 7 | 6 | 5 | 4 | 4 |
| $k_1 = 2.2$ | | | | | | | |
| $m_u$ | 6 | 5 | 4 | 4 | 3 | 3 | 3 |
| $m_p$ | 11 | 7 | 6 | 5 | 4 | 3 | 3 |

TABLE 6. The Minimum Number of Subsamples, $m$, Needed for Power $\geq 0.80$, When $k_0 = 1.5$ and $\alpha = 0.01$

| | $n=4$ | $n=5$ | $n=6$ | $n=7$ | $n=8$ | $n=9$ | $n=10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 1.9$ | | | | | | | |
| $m_u$ | 25 | 20 | 17 | 14 | 13 | 11 | 10 |
| $m_p$ | >100 | 52 | 32 | 23 | 18 | 15 | 12 |
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 17 | 14 | 12 | 10 | 9 | 8 | 7 |
| $m_p$ | 52 | 28 | 19 | 14 | 12 | 10 | 9 |
| $k_1 = 2.1$ | | | | | | | |
| $m_u$ | 13 | 10 | 9 | 8 | 7 | 6 | 5 |
| $m_p$ | 31 | 18 | 13 | 10 | 8 | 7 | 6 |
| $k_1 = 2.2$ | | | | | | | |
| $m_u$ | 10 | 8 | 7 | 6 | 5 | 5 | 4 |
| $m_p$ | 21 | 13 | 10 | 8 | 6 | 6 | 5 |

TABLE 7. The Minimum Number of Subsamples, $m$, Needed for Power $\geq 0.80$, When $k_0 = 1.6$ and $\alpha = 0.10$

| | $n = 4$ | $n = 5$ | $n = 6$ | $n = 7$ | $n = 8$ | $n = 9$ | $n = 10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 12 | 10 | 8 | 7 | 6 | 6 | 5 |
| $m_p$ | 53 | 23 | 14 | 11 | 8 | 7 | 6 |
| $k_1 = 2.1$ | | | | | | | |
| $m_u$ | 9 | 7 | 6 | 5 | 5 | 4 | 4 |
| $m_p$ | 22 | 12 | 8 | 7 | 6 | 5 | 4 |
| $k_1 = 2.2$ | | | | | | | |
| $m_u$ | 6 | 5 | 4 | 4 | 3 | 3 | 3 |
| $m_p$ | 13 | 8 | 6 | 5 | 4 | 3 | 3 |
| $k_1 = 2.3$ | | | | | | | |
| $m_u$ | 5 | 4 | 4 | 3 | 3 | 3 | 2 |
| $m_p$ | 9 | 6 | 4 | 4 | 3 | 3 | 3 |

TABLE 8. The Minimum Number of Subsamples, $m$, Needed for Power $\geq 0.80$, When $k_0 = 1.6$ and $\alpha = 0.05$

| | $n = 4$ | $n = 5$ | $n = 6$ | $n = 7$ | $n = 8$ | $n = 9$ | $n = 10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 17 | 14 | 11 | 10 | 9 | 8 | 7 |
| $m_p$ | 83 | 35 | 21 | 16 | 12 | 10 | 9 |
| $k_1 = 2.1$ | | | | | | | |
| $m_u$ | 12 | 9 | 8 | 7 | 6 | 5 | 5 |
| $m_p$ | 34 | 18 | 12 | 9 | 8 | 7 | 6 |
| $k_1 = 2.2$ | | | | | | | |
| $m_u$ | 9 | 7 | 6 | 5 | 5 | 4 | 4 |
| $m_p$ | 20 | 12 | 8 | 7 | 6 | 5 | 4 |
| $k_1 = 2.3$ | | | | | | | |
| $m_u$ | 7 | 6 | 5 | 4 | 4 | 3 | 3 |
| $m_p$ | 13 | 8 | 6 | 5 | 4 | 4 | 3 |

TABLE 9. The Minimum Number of Subsamples, $m$, Needed for Power $\geq 0.80$, When $k_0 = 1.6$ and $\alpha = 0.01$

| | $n = 4$ | $n = 5$ | $n = 6$ | $n = 7$ | $n = 8$ | $n = 9$ | $n = 10$ |
|---|---|---|---|---|---|---|---|
| $k_1 = 2.0$ | | | | | | | |
| $m_u$ | 28 | 22 | 19 | 16 | 14 | 13 | 11 |
| $m_p$ | >100 | 64 | 39 | 27 | 21 | 17 | 15 |
| $k_1 = 2.1$ | | | | | | | |
| $m_u$ | 19 | 15 | 13 | 11 | 10 | 9 | 8 |
| $m_p$ | 64 | 33 | 22 | 17 | 13 | 11 | 10 |
| $k_1 = 2.2$ | | | | | | | |
| $m_u$ | 14 | 12 | 10 | 8 | 7 | 7 | 5 |
| $m_p$ | 36 | 21 | 15 | 11 | 9 | 8 | 7 |
| $k_1 = 2.3$ | | | | | | | |
| $m_u$ | 11 | 9 | 8 | 7 | 5 | 5 | 4 |
| $m_p$ | 24 | 15 | 11 | 9 | 7 | 6 | 5 |

As seen in the tables, for the threshold power of 0.80 the required number of subsamples is always smaller when using the un-pooled variance estimator. Consequently, it is recommended that, when possible, the un-pooled variance estimator be used to assess the capability of a process. This recommendation is consistent with and supportive of the conclusion of Cryer and Ryan (1990) that the "overall" or "long-term" variance estimator is superior to a control-chart related "within" or "short-term" variance estimator. Furthermore, when the process has undergone a change in variation, the un-pooled estimator captures all of the variation, whereas the pooled captures only the within-subsample component.

We also want to draw attention to the fact that the critical values differ substantially for the pooled and un-pooled cases for fixed $\alpha$, $m$, and $n$. Hence, it is important to use the correct critical value that corresponds to the estimator used. For examples illustrating this, see Vännman and Hubele (2001, 2003).

In all discussions concerning power, there are the considerations of significance level and sample size. In this application, the sample size $n$ and the number of subsamples $m$ affect the power. We recommend specifying a minimum power of 0.80, which we believe will provide the user with a reasonable recommended number of subsamples. The larger the subsamples, the fewer the number of subsamples needed for all cases investigated.

If, instead, the user requires a fixed number of subsamples, the un-pooled estimator always provides more power than the pooled estimator. However, as the subsample size increases and as the difference between the null and alternative hypothesized values, $k_0$ and $k_1$, increases, the power becomes nearly the same for both estimators. However, the critical values differ, as noted earlier.

If it is of interest to use other threshold values for $k_0$, $k_1$, and the power than those presented here, corresponding results are easily obtained using the formulas provided above. For theoretical foundations of an estimator of $C_{pk}$ when a pooled variance estimator is used, see Vännman and Hubele (2003).

The sub-sample size recommendations made in this paper are based on power computations in hypothesis testing. After having performed the test and reached a conclusion about having a capable process, it is always a good idea to construct a one-sided lower confidence interval. Zimmer et al. (2001) and Boyles (1991) suggested approximate confidence intervals for $C_{pm}$ when considering the single sample case. These intervals are based on the large sample approximation of the non-central $\chi^2$-distribution. Furthermore, they either assume the non-centrality parameter $\lambda$ to be known or to be estimated, without taking into consideration how the randomness in the estimator of $\lambda$ affects the confidence level. By using an estimator of $\lambda$, yet another approximation, apart from the large sample approximation, is introduced. Similar approximate confidence intervals can be calculated, when using sub-samples, by utilizing Equations (13) or (20). However, there is still the unanswered question about how the number of sub-samples will affect the approximations of the confidence level for the un-pooled and pooled case, respectively.

## Appendix A

In Figures 2-4 we can see quite different behaviors of the power graphs, as functions of $\delta$ depending on which estimator is used. The reason for these different behaviors is the way in which the two distributions given in Equations (23) and (28) and the critical values given in Equations (26) and (30) depend on $\delta$, $m$, and $n$. To help explain these differences, we use the examples shown in Figures 2b, 3b, and 4b. We note that $m$ and $n$ in these examples have been chosen to illustrate the use of the tables, where for fixed value of $n = 4$ the minimum number of subsamples needed to have a power of at least 0.80 for both the un-pooled and the pooled case are given. This implies that $m$ will differ between the pooled and un-pooled case. Here, we explain the difference in behaviors seen in Figures 2-4 without going into details by comparing the behavior of the power in general. Power comparisons when $m$ and $n$ are kept equal in the pooled and un-pooled case can be found in work by Vännman and Hubele (2003). We now consider the following two situations, the power graphs of which are shown in Figures 2-4.

### (i) The Pooled Case

The null hypothesis is $H_0 : C_{pm} = k_0 = 4/3$, the alternative hypothesis is $H_1 : C_{pm} = k_1 = 1.9$, the significance level $\alpha = 0.05$, $n = 4$, and $m = 14$. We use the pooled variance estimator, i.e., we use $\widehat{C}_{pm,p}$ as the estimator of $C_{pm}$. Then, we have the critical value $c_{0.05,p} = 1.8540$ according to Equation (26). The power is, according to Equation (27),

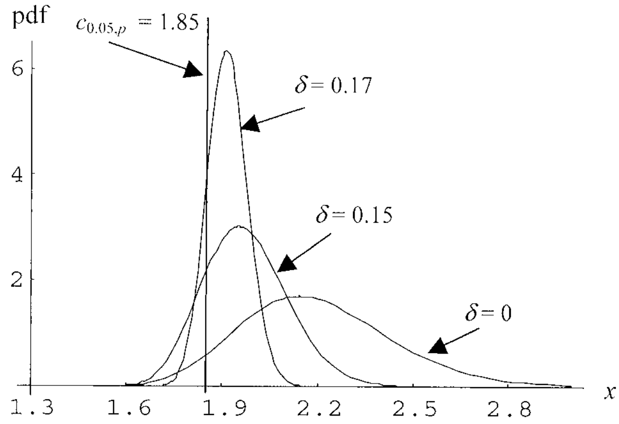


FIGURE 5. Situation (i), the Pooled Case with $m = 14$, $n = 4$, $\alpha = 0.05$. The Probability Density Functions of $\widehat{C}_{pm,p}$, Given that $C_{pm} = 1.9$, when $\delta = 0,\ 0.15,\ 0.17$. The Area to the Right of $c_{0.05,p} = 1.85$ Under the Corresponding Probability Density Function Shows the Power for $C_{pm} = 1.9$. For $\delta = 0$ the Power is 0.95, for $\delta = 0.15$ it is 0.82, and for $\delta = 0.17$ it is 0.84.

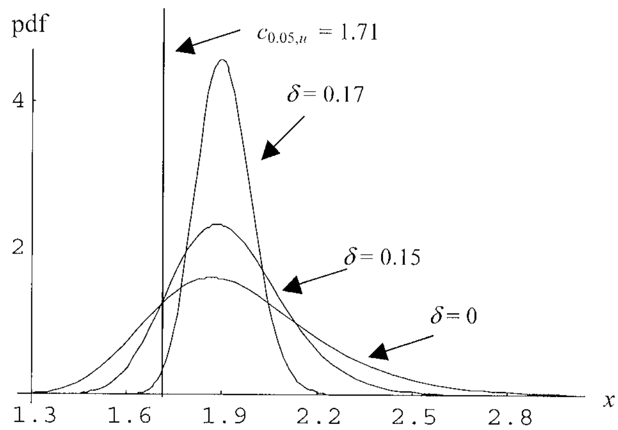


FIGURE 6. Situation (ii), the Un-Pooled Case with $m = 7$, $n = 4$, $\alpha = 0.05$. The Probability Density Functions of $\widehat{C}_{pm,u}$, Given that $C_{pm} = 1.9$, when $\delta = 0,\ 0.15,\ 0.17$. The Area to the Right of $c_{0.05,u} = 1.85$ Under the Corresponding Probability Density Function Shows the Power for $C_{pm} = 1.9$. For $\delta = 0$ the Power is 0.81, for $\delta = 0.15$ it is 0.89, and for $\delta = 0.17$ it is 0.99.

$$\begin{aligned}\mathrm{Pow}_p(\delta) &= \mathrm{Pow}_p(\delta, 14, 4, 0.05, 4/3, 1.9)\\ &= F_\zeta\left(\frac{58.8133}{1-32.49\delta^2}\right),\end{aligned} \tag{A1}$$

where $\zeta$ is a non-central $\chi^2$-distributed random variable with 43 degrees of freedom and non-centrality parameter

$$\lambda = \lambda_p(\delta) = \frac{1819.44\delta^2}{1-32.49\delta^2}. \tag{A2}$$

The way the power, $\mathrm{Pow}_p(\delta)$, depends on $\delta = (\mu - T)/d$ is quite complex, which might be easier to understand using Patnaik's approximation for the power. From Equation (27) we obtain

$$\mathrm{Pow}_p(\delta) \approx F_{\chi^2_{f_p}}\left(\frac{58.8133}{g_p(1-32.49\delta^2)}\right).$$

Both the degrees of freedom $f_p$ in the central $\chi^2$-distribution and the constant $g_p$ used in Patnaik's approximation depend on $\delta$. From Equations (10) and (11) we have, since $mn - (m-1) = 56 - 13 = 43$,

$$f_p = \frac{(43+\lambda_p^2(\delta))^2}{43+2\lambda_p(\delta)} \quad \text{and} \quad g_p = \frac{43+2\lambda_p(\delta)}{43+\lambda_p(\delta)},$$

where $\lambda_p(\delta)$ is given in Equation (A2).

Plotting $\mathrm{Pow}_p(\delta)$ as a function of $\delta$ we obtain the result shown in Figure 2b, where we can see that the power given $C_{pm} = k_1 = 1.9$ is at least 0.80 for all values of $\delta$.

### (ii) The Un-pooled Case

The conditions of the hypotheses are the same except $m = 7$. We use the un-pooled variance estimator, i.e., we use $\widehat{C}_{pm,u}$ as the estimator of $C_{pm}$. Then we have the critical value $c_{0.05,u} = 1.7148$ according to Equation (30). The power is, according to Equation (31),

$$\begin{aligned}\mathrm{Pow}_u(\delta) &= \mathrm{Pow}_u(\delta, 7, 4, 0.05, 4/3, 1.9)\\ &= F_\xi\left(\frac{34.3765}{1-32.49\delta^2}\right),\end{aligned} \tag{A3}$$

where $\xi$ is a non-central $\chi^2$-distributed random variable with 28 degrees of freedom and non-centrality parameter

$$\lambda = \lambda_u(\delta) = \frac{909.72\delta^2}{1-32.49\delta^2}. \tag{A4}$$

Using Patnaik's approximation in Equation (31) we obtain

$$\mathrm{Pow}_u(\delta) \approx F_{\chi^2_{f_u}}\left(\frac{909.72}{g_u(1-32.49\delta^2)}\right),$$

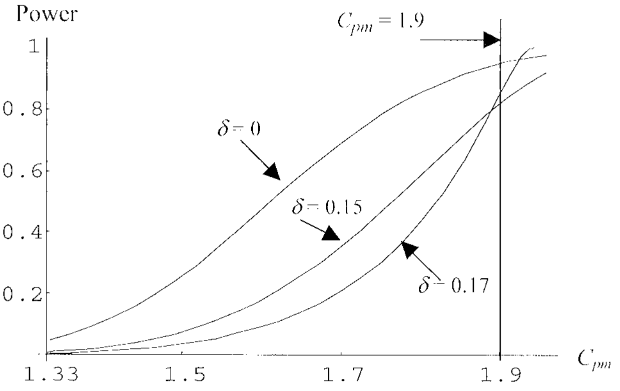


FIGURE 7. The Power as a Function of $C_{pm}$ for $\delta = 0$, 0.15, 0.17 When Testing $H_0 : C_{pm} = k_0 = 4/3$ in Situation (i), the Pooled Case with $m = 14$, $n = 4$, $\alpha = 0.05$.

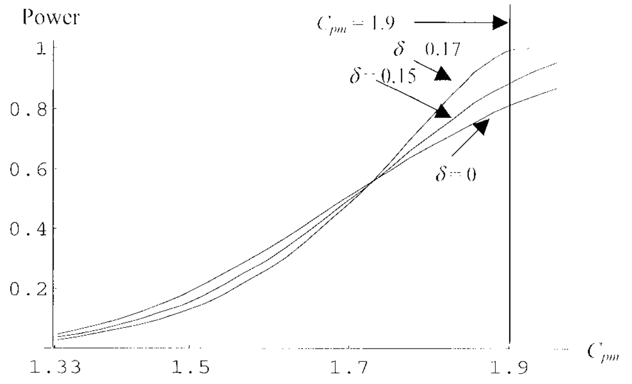


FIGURE 8. The Power as a Function of $C_{pm}$ for $\delta = 0$, 0.15, 0.17 When Testing $H_0 : C_{pm} = k_0 = 4/3$ in Situation (ii), the Un-Pooled Case with $m = 7$, $n = 4$, $\alpha = 0.05$.

where, since $mn = 28$,

$$f_u = \frac{(28 + \lambda_u^2(\delta))^2}{28 + 2\lambda_u(\delta)} \quad \text{and} \quad g_p = \frac{28 + 2\lambda_u(\delta)}{28 + \lambda_u(\delta)},$$

with $\lambda_u(\delta)$ given in Equation (A4).

Plotting $\mathrm{Pow}_p(\delta)$ as a function of $\delta$ we obtain the result shown in Figure 3b. The power, given that $C_{pm} = k_1 = 1.9$, is at least 0.80 for all values of $\delta$ plotted.

When we compare the two expressions in Equations (A1) and (A3) used to calculate the two power curves $\mathrm{Pow}_p(\delta)$ and $\mathrm{Pow}_u(\delta)$ above, we see that they differ in three aspects although they are both based on non-central $\chi^2$-distributions. They have different degrees of freedom as well as different non-centrality parameters. Furthermore, the arguments in Equations (A1) and (A3) differ as a result of different critical values.

To better understand how the two different power expressions behave when $\delta$ changes in the interval [0, 0.175], we plot the probability density functions for $\widehat{C}_{pm,p}$ and $\widehat{C}_{pm,u}$ in Figure 5 and Figure 6 for $\delta = 0$, 0.15, and 0.17. The critical values $c_{0.05,p} = 1.85$ and $c_{0.05,u} = 1.71$, respectively, have been included in the figures to facilitate the interpretation of the power, which is the area to the right of the critical value under the probability density function.

From Figure 5 we can see that, for the pooled case, the location of the distribution depends strongly on $\delta$. When $\delta$ is small the probability density function is shifted with its center to the right of 1.9, which is the true value of $C_{pm}$. When $\delta$ increases the probability density function is shifted to the left and centers more closely to 1.9. This shift of the distribution to the left, when $\delta$ increases, explains to some extent the decrease of the power for $\delta$-values in the interval [0, 1.5] shown in Figure 2b. We can also see that when $\delta$ increases from 0 to 0.175 the spread of the distribution decreases, and when $\delta$ is close to the border value 0.175 the probability density function gets very narrow. We also see that the distribution becomes less skewed when $\delta$ increases. All these changes in location, spread, and skewness taken together help to explain the behavior of the power for a given value of $C_{pm}$ as a function of $\delta$.

From Figure 6 we can see that, for the un-pooled case, the probability density function is centered in the vicinity of 1.9 for all values of $\delta$ and the location does not shift very much when $\delta$ increases. Furthermore, the distributions are less skewed for the un-pooled case compared to the pooled case. In Figure 6 we can see as well as in Figure 5 that the spread decreases with increasing $\delta$-values. This is the main reason that the power in Figure 3b increases when $\delta$ increases.

In addition, we note that the critical value for the un-pooled case is smaller than the critical value for the pooled case. Since the power is the area under the probability density function to the right of the critical value, this fact also affects the power.

To demonstrate yet another aspect of how the power depends on $\delta$, the power as a function of $C_{pm}$ has been plotted for $\delta = 0$, 0.15, and 0.17 for situation (i) and (ii), respectively, in Figures 7-8. Here, we can see that in each of the two situations different values of $\delta$ give rise to different power curves as functions of $C_{pm}$. We note that the number of subsamples, $m$, differs between Figure 7 and Figure 8 due to reasons explained in the beginning of the appendix. In Figures 7 and 8, the value when $C_{pm} = 1.9$ is marked, whereas Figures 2b, 3b, and 4b show the power for this value of $C_{pm}$ when $\delta$ varies.

## Appendix B

To derive the result in Equation (32) for choosing $m_u$, we consider $\text{Pow}(\delta; m, n, \alpha, k_0, k_1)$ in Equation (31) as a function of $\delta$. Through numerical investigations we find that $\text{Pow}(\delta)$ has its minimum when $\delta = 0$. For examples see Figures 3 and 4. Arizono et al. (1997) proved a similar result using an approximation to the percentile of the $\chi^2$-distribution when investigating the quadratic loss function. Hence, using Equations (29)-(31) we find

$$\text{Pow}_u(\delta; m, n, \alpha, k_0, k_1) \geq \text{Pow}_u(0; m, n, \alpha, k_0, k_1) = F_\xi\left(\frac{k_1^2}{k_0^2} \cdot \chi^2_{\alpha,mn}\right),$$

where $\xi$ is distributed according to a central $\chi^2$-distribution with $mn$ degrees of freedom. This probability will be at least $1 - \beta$ if

$$\frac{k_1^2}{k_0^2} \geq \frac{\chi^2_{1-\beta,mn}}{\chi^2_{\alpha,mn}},$$

where $\chi^2_{\alpha,mn}$ represents the $\alpha$th quantile from the $\chi^2$-distribution with $mn$ degrees of freedom. Using the Wilson-Hilferty approximation for a quantile from the $\chi^2$-distribution as given by Johnson, Kotz, and Balakrishnan (1994, p. 427) we arrive at the result in Equation (32) after straightforward, but tedious, calculations.

## Acknowledgments

The authors are grateful for the financial support provided by the National Science Foundation, Grant No. DMI-9700581, and the Swedish Research Council, project number 621-2002-3823. In addition, we appreciate the very thorough and thoughtful comments of two anonymous referees that helped to improve and clarify our paper.

## References

ARIZONO, I.; KANAGAWA, A.; OHTA, H.; WATAKABE, K.; and TATEISHI, K. (1997). "Variable Sampling Plans for Normal Distribution Indexed by Taguchi's Loss Function". *Naval Research Logistics* 44, pp. 591–603.

BOYLES, R. A. (1991). "The Taguchi Capability Index". *Journal of Quality Technology* 23, pp. 17–26.

CHAN, L. K.; CHENG, S. W.; and SPIRING, F. A. (1988). "A New Measure of Process Capability: $C_{pm}$". *Journal of Quality Technology* 20, pp. 162–175.

CHENG, S. W. (1992). "Is the Process Capable? Tables and Graphs in Assessing $C_{pm}$". *Quality Engineering* 4, pp. 563–576.

CRYER, J. D. and RYAN, T. P. (1990). "The Estimation of Sigma for an X Chart: $\overline{MR}/d_2$ or $S/c_4$?" *Journal of Quality Technology* 22, pp. 187–192.

DELERYD, M. and VÄNNMAN, K. (1999). "Process Capability Plots—A Quality Improvement Tool". *Quality and Reliability Engineering* 15, pp. 1–15.

JOHNSON, N. L.; KOTZ, S.; and BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions Volume 1*, 2nd ed. Wiley, New York, NY.

JOHNSON, N. L.; KOTZ, S.; and BALAKRISHNAN, N. (1995). *Continuous Univariate Distributions Volume 2*, 2nd ed. Wiley, New York, NY.

KIRMANI, S. N. U. A.; KOCHERLAKOTA, K.; and KOCHERLAKOTA, S. (1991). "Estimation of $\sigma$ and the Process Capability Index Based on Subsamples". *Communications in Statistics—Theory and Methods* 20, pp. 275–291.

KOTZ, S. and JOHNSON, N. L. (1993). *Process Capability Indices*. Chapman and Hall, London, UK.

KOTZ, S. and JOHNSON, N. L. (2002). "Process Capability Indices—A Review, 1992-2000 (with discussion)". *Journal of Quality Technology* 34, pp. 2–53.

KOTZ, S. and LOVELACE, C. R. (1998). *Introduction to Process Capability Indices: Theory and Practice*. Arnold, London, UK.

LI, H.; OWEN, D. B.; and BORREGO, S. A. (1990). "Lower Confidence Limits on Process Capability Indices Based on the Range". *Communications in Statistics—Simulation* 19, pp. 1–24.

RODRIGUEZ, R. "Discussion of 'Process Capability Indices—A Review, 1992–2000' ". *Journal of Quality Technology* 34, pp. 28–31.

VÄNNMAN, K. (1997). "Distribution and Moments in Simplified Form for a General Class of Capability Indices". *Communications in Statistics—Theory and Methods* 26, pp. 159–179.

VÄNNMAN, K. and HUBELE, N. F. (2001). "The Distribution of $C_{pm}$ When Using a Pooled Variance Estimator with Some Results". *American Statistical Association 2001 Proceedings of the Section on Quality and Productivity* on CD-ROM.

VÄNNMAN, K. and HUBELE, N. F. (2003). "Distributional Properties of Estimated Capability Indices Based on Subsamples". *Quality and Reliability Engineering International* 19, pp. 111–128.

ZIMMER, L.; HUBELE, N. F.; and ZIMMER, W. J. (2001) "Confidence Intervals and Sample Size Determination for $C_{pm}$". *Quality and Reliability Engineering International* 17, pp. 51–68.