Quality Engineering

ISSN: 0898-2112 (Print) 1532-4222 (Online) Journal homepage: http://www.tandfonline.com/loi/lqen20

A short note on the effect of sample size on the estimation error in C_p

Zahra Sedighi Maman, W. Wade Murphy, Saeed Maghsoodloo, Fatemah Haji Ahmadi & Fadel M. Megahed

QUALITY ENGINEERING
http://dx.doi.org/10.1080/08982112.2016.1172091

Taylor & Francis
Taylor & Francis Group

A short note on the effect of sample size on the estimation error in C_p

Zahra Sedighi Maman[a], W. Wade Murphy[a], Saeed Maghsoodloo[a], Fatemah Haji Ahmadi[b], and Fadel M. Megahed[a]

[a]Department of Industrial and Systems Engineering, Auburn University, Auburn, Alabama; [b]Department of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran

ABSTRACT
Process capability indices such as C_p are used extensively in manufacturing industries to assess processes in order to decide about purchasing. In practice, the parameter for calculating C_p is rarely known and is frequently replaced with estimates from an in-control reference sample. This article explores the optimal sample size required to achieve a desired error of estimation using absolute percentage error of different C_p estimates. Moreover, some practical tools are created to allow practitioners to find sample size in different situations.

KEYWORDS
absolute percentage error; Phase I; process capability; six sigma; standard deviation

Introduction and literature review

Evaluating the capability of a manufacturing process is an important concept that has received much interest in six sigma, lean manufacturing and statistical process control (see Kotz and Johnson 1993; Kumar et al. 2006; Chen et al. 2010). In general, process capability compares the output of an *in-control* and steady process to the preset engineering specification limits by using *capability indices*. For example, the most popular capability index (C_p) forms the "ratio of the spread between the process specifications (the specification "width") to the spread of the process values, as measured by 6 process standard deviation units (the process "width")" (see NIST/SEMATECH e-Handbook of Statistical Methods 2012). Mathematically, C_p is defined as:

$$C_p = (\text{USL} - \text{LSL})\,/6\sigma, \qquad [1]$$

where USL, LSL, and σ are the upper specification limit, lower specification limit, the process standard deviation, respectively. Based on Eq. [1] and under the assumptions that the process is centered, C_p can be used to easily quantify the % rejects of a process (see Table 1).

The calculations presented in Table 1 assume that the process standard deviation is known. In practice, however, process parameters such as σ are rarely known, and they are estimated based on a suitable baseline sample. There are two possibilities for getting such estimates: (1) based on a Phase I control chart; or (2) from a dedicated baseline sample. For our purposes, these two scenarios are identical since the effect of estimation error on C_p is solely based on sample size. When σ is to be estimated, process capability would be defined as:

$$\widehat{C_p} = (\text{USL} - \text{LSL})\,/6\hat{\sigma}, \qquad [2]$$

where $\hat{\sigma}$ represents any appropriate estimator for σ. Note that any error in estimating σ would result in an incorrect estimator of the true process capability. The problem of estimating σ can be divided into two parts: (a) what is the best (robust, unbiased, and/or min variance) estimator for σ? (b) What sample size is needed such that the effect of estimation error can be neglected? The selection of the best estimator for sigma in part *a* was considered by several researchers including Kirmani, Kocherlakota, and Kocherlakota (1991), Derman and Ross (1995), Ravindra Khattree (1999), and Mahmoud et al. (2010). In the context of the control charting literature, the effect of parameter estimation on a control chart's properties has been reviewed by Jensen et al. (2006) and Jones-Farmer et al. (2014). These articles show that estimated parameters can have a significant effect on both the *in-control* and *out-of-control* performance of control charts, especially with small to moderate sample sizes.

CONTACT Fadel M. Megahed fmegahed@auburn.edu Department of Industrial and Systems Engineering, Auburn University, Auburn, AL 36830.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lqen.

Table 1. Practical use of process capability indices.

USL—LSL	6σ	8σ	10σ	12σ
Cp	1.00	1.33	1.66	2
Defect Rates	0.27%	64 parts per million	0.6 parts per million	2 parts per billion
% of spec used	100	75	60	50

For part *b*, Franklin (1999), Zimmer, Hubele, and Zimmer (2001), Pearn and Ming-Hung (2003), and Wua and Kuo (2004) have investigated the sample sizes needed by considering a lower confidence interval approach as a basis for the decision. They have assumed that the process is *in-control* (or steady) and its output is normally distributed (or an appropriate transformation can be applied to not violate the normality assumption). It should be noted that these articles use the ratio of actual process capability over the estimated one ($\frac{C_p}{\hat{C}_p}$) and offer sample size recommendation based on the associated confidence interval. Unfortunately, this approach is somewhat limiting in practice. We highlight three potential issues in using the ratio ($\frac{C_p}{\hat{C}_p}$) as a basis for decision-making.

(a) The selection of a suitable value for this ratio is not clear. For example, Franklin (1999) and Wu and Kuo (2004) have used arbitrary ratios such as 0.8, 0.85, and/or 0.9 to determine the sample size. There are no justifications for how a practitioner should select such a ratio for determining the sample size (with the exception that ratios closer to 1 are preferred).

(b) The use of this ratio might lead practitioners to believe that the difference from 1 is a measure of the estimation error. For example, one might think that using a ratio of 0.8 would imply that the error is 0.2. However, the associated absolute percentage error (APE) in this case is 0.25 which is larger than the 0.2 (incorrect interpretation of error). While this might seem intuitive to a statistics-*savvy* audience, this is not necessarily the case for practitioners who use process capability indices (instead of calculating the probability of being outside the spec. limits).

(c) From a statistical perspective, the ratio of ($\frac{C_p}{\hat{C}_p}$) is a random variable (r.v.), i.e., it has a different expected value and variance for different sample sizes (and/or different values of $\hat{\sigma}$). By pre-specifying a single value, one would discard the stochastic nature that is associated with estimating C_p.

To address these three issues, we consider the use of APE (see "Methodology" for a mathematical definition) to determine the sample size needed such that the effect of estimation on the process capability index can be neglected (based on thresholds that are set by the practitioner). In this article, we consider single and multiple sampling procedures. By using APE in the decision-making process, we can achieve the following.

(a) The interpretation of APE is much simpler than the aforementioned ratio. Specifically, the use of APE allows us to consider the estimation error as a function of sample size.

(b) We consider APE as a random variable as detailed later in this article. This allows us to consider both the expected value and the standard deviation of APE when calculating the sample size. It is important to note that the ability to calculate the standard deviation of the APE allows us to consider the between samples variation (each typically considering one baseline sample) in estimating C_p.

It should be noted that we consider the calculation of sample size in the case of single and multiple sampling procedures, as well as through using different estimators of σ. To ensure the broad reach of this approach, we provide a toolkit to allow practitioners to find appropriate sample size based on simple criteria (see the Appendix for more details).

In this article, the process is supposed to be *in-control*, centered and the quality characteristic follows a normal distribution. The next section presents definitions for the expected value of APE and its standard deviation, and the procedure for calculating the sample size based on the different estimators for σ. In the "Results and discussion" section, we present numerical results to highlight how our approach can be used in practice. Finally, in the last section, we offer some concluding remarks. The mathematical derivations, codes used, and an overview of the practitioner toolkit are provided in Appendices A, B, and C, respectively.

Methodology

The single sampling case

In this section, we consider scenarios where practitioners attempt to estimate the process standard deviation based on a single baseline sample. In particular, two different estimators for σ are discussed; s and s/c_4. Below, we show how practitioners can use the APE statistic to determine the sample size needed such that the effect of any estimation error on C_p can be neglected.

s as an estimator for σ

The sample standard deviation, s, is widely used to estimate the population/process standard deviation. Let $i = 1, 2, \ldots$ be an independent sample of size n drawn from a process that is normally distributed with constant, but unknown parameters (μ, σ^2). Then, s can be calculated as:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}, \quad [3]$$

where $\bar{x}$ is the sample mean. When s is used to estimate σ, Eq. [2] can be re-written as:

$$\widehat{Cp_s} = (\text{USL} - \text{LSL})/6s. \quad [4]$$

Note that we added the subscript s to denote that the use of the sample standard deviation. In this case, the *Absolute Percentage Error* for C_p can be defined as:

$$\text{APE}_s = \left| \frac{Cp - \widehat{Cp_s}}{Cp} \right| = \left| \frac{\frac{\text{USL}-\text{LSL}}{6\sigma} - \frac{\text{USL}-\text{LSL}}{6s}}{\frac{\text{USL}-\text{LSL}}{6\sigma}} \right| = \left| 1 - \frac{\sigma}{s} \right|. \quad [5]$$

Since APE_s is a r.v., we can define its expected value and standard deviation based on $\text{U} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$. Then the expected value and standard deviation of the APE_s can be formulated as:

$$\text{APE}_s = \text{f}(\text{U}) = \left| 1 - \left(\frac{\text{U}}{n-1} \right)^{-\frac{1}{2}} \right|, \quad [6]$$

$$\text{E}(\text{APE}_s) = \int_0^\infty \text{f}(\text{U})\, \text{g}(\text{U})\, \text{du}, \text{ and} \quad [7]$$

$$\text{SD}(\text{APE}_s) = \left(\text{E}\left(\text{APE}_s{}^2\right) - \text{E}^2(\text{APE}_s) \right)^{\frac{1}{2}}. \quad [8]$$

Note that g(U) is the probability density function for U based on the χ^2_{n-1} distribution. Based on the Eqs. [6]–[8], we can define a function that allows a practitioner to find an appropriate sample size based on a pre-specified error criterion as shown in Eq. [9]:

$$\text{P}(\text{APE}_s < \mathit{Max\ APE}) > 1 - \alpha, \quad [9]$$

where $1-\alpha$ represents a confidence level such that $0 < \alpha < 1$. Note that the choice of α represents the risk threshold that the practitioner is willing to take. For example, $\alpha = 0.05$ would mean that in 95% of the samples APE_s will be smaller than the Max APE. In Eq. [9], for a given Max APE (representing a user's required level of accuracy) and a pre-defined α value, n is the only unknown. Thus, Eq. [9] can be used to obtain the smallest sample size that meets these two criteria. The section titled "Results for the single sampling scenarios" provides numerical solutions for different combinations of Max APE and α.

s/c4 as an estimator for σ

It is well documented that s is a biased estimator for the population standard deviation, and thus, a correction factor c_4 is often used to eliminate the bias (see, e.g., Montgomery Runger, and Hubele 2009; Mahmoud et al. 2010). Note c_4 is a function of n:

$$c_4 = c_4(n) = \left(\frac{2}{n-1} \right)^{\frac{1}{2}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}. \quad [10]$$

In this situation, $\hat{\sigma} = s/c_4$ and $\widehat{Cp_{\frac{s}{c_4}}} = \frac{\text{USL}-\text{LSL}}{6\frac{s}{c_4}}$. By substituting s by s/c_4 in Eqs. [5]–[9], one could easily obtain an expression where $(\text{APE}_{\frac{s}{c_4}} < \text{Max APE}) > 1 - \alpha$. This expression can then be used to obtain the smallest sample size that meets Max APE and a pre-defined α value given that s/c_4 is used to estimate the process standard deviation. For the sake of completion, we provide the detailed mathematical expressions in the Appendix.

Using multiple samples for estimating the process standard deviation

Similar to the previous section "The single sampling case," we provide details for two commonly used estimators for σ when multiple samples are used. The details for using these estimators are provided below.

Using the pooled sample standard deviation (Sp)

Let x_{ij} ($i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$) be m independent baseline samples of size n drawn from a process that is normally distributed with constant, but unknown parameters (μ, σ^2). The pooled sample standard deviation, S_p, can be used to estimate the process standard deviation:

$$S_p = \left(\frac{1}{m} \sum_{i=1}^{m} s_i^2 \right)^{\frac{1}{2}} = \left(\frac{1}{m(n-1)} \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 \right)^{\frac{1}{2}}. \quad [11]$$

Consequently, C_p and the APE_{S_p} can be calculated as:

$$\widehat{Cp_{S_p}} = (\mathrm{USL} - \mathrm{LSL})/6S_p \text{, and} \quad [12]$$

$$\mathrm{APE}_{S_p} = \left|\frac{Cp - \widehat{Cp_{S_p}}}{Cp}\right| = \left|\frac{\frac{\mathrm{USL}-\mathrm{LSL}}{6\sigma} - \frac{\mathrm{USL}-\mathrm{LSL}}{6\,S_p}}{\frac{\mathrm{USL}-\mathrm{LSL}}{6\sigma}}\right| = \left|1 - \frac{\sigma}{S_p}\right|. \quad [13]$$

Expressions for expected value and standard deviation of APE_{S_p} can be derived by replacing s by S_p in Eqs. [6]–[8] and using $\mathrm{U} = \frac{m(n-1)S_p^2}{\sigma^2} \sim \chi^2{}_{m(n-1)}$. Based on this information, we can now define a function that allows us to calculate the required sample size for a given α and Max APE:

$$\mathrm{P}\left(\mathrm{APE}_{S_p} < \text{Max APE}\right) > 1 - \alpha. \quad [14]$$

Using $\frac{\bar{s}}{c_4}$ for estimating the process standard deviation

Another approach for estimating σ when multiple baseline samples are drawn can be obtained by using the estimator $\hat{\sigma} = \bar{s}/c_4$, which can be seen as the multiple sample extension for the method highlighted in the section titled "s/c4 as an estimator for $\boldsymbol{\sigma}$". Similar to our discussion for S_p, let x_{ij} ($i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$) be m independent samples of size n from a $N \sim (\mu,\ \sigma^2)$ process. $\bar{s}$ is defined as:

$$\bar{s} = \frac{1}{m}(s_1 + s_2 + \cdots + s_m), \quad [15]$$

where s_i is the standard deviation for sample i, which can be calculated by Eq. [3]. By replacing s by $\bar{s}/c_4$ in Eqs. [4] and [5], we can obtain the following:

$$\widehat{Cp_{\frac{\bar{s}}{c_4}}} = \frac{\mathrm{USL} - \mathrm{LSL}}{6\,\frac{\bar{s}}{c_4}}, \text{ and} \quad [16]$$

and

$$\mathrm{APE}_{\frac{\bar{s}}{c_4}} = \left|\frac{Cp - \widehat{Cp_{\frac{\bar{s}}{c_4}}}}{Cp}\right| = \left|\frac{\frac{\mathrm{USL}-\mathrm{LSL}}{6\sigma} - \frac{\mathrm{USL}-\mathrm{LSL}}{6\,\frac{\bar{s}}{c_4}}}{\frac{\mathrm{USL}-\mathrm{LSL}}{6\sigma}}\right| = \left|1 - c_4\frac{\sigma}{\bar{s}}\right|. \quad [17]$$

Similar to our discussion in "The single sampling case," $\mathrm{APE}_{\frac{\bar{s}}{c_4}}$ is a random variable. To calculate its expected value and standard deviation, let $\mathrm{Q} = \frac{\hat{\sigma}}{\sigma}$ such that the distribution of Q is independent of σ. Based on the derivations in Patnaik (1950) and Chen (1998), Q is a scaled chi random variable $\frac{c\chi_v}{\sqrt{v}}$, and the probability density of $\frac{c\chi_v}{\sqrt{v}}$ is:

$$\mathrm{g}(q, v, c) = \frac{2}{c}\frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)}\left(\frac{q}{c}\right)^{v-1}\exp\left(-\frac{v}{2}\left(\frac{q}{c}\right)^2\right), \quad [18]$$

where $v = (-2 + 2\sqrt{1+2t})^{-1}$ and $c = 1 + \frac{1}{4v} + \frac{1}{32v^2} - \frac{5}{128v^3}$. To calculate v and c, we follow the approach of Patnaik (1950) and Chen (1998) who used $r = (-2 + 2\sqrt{1+2M_1})^{-1}$, $t = M_1 + \frac{1}{16r^3}$, and $M_1 = \frac{1-c_4(n)^2}{mc_4(n)^2}$. Therefore, the APE, expected value and standard deviation can be obtained as:

$$\mathrm{APE}_{\frac{\bar{s}}{c_4}} = f(\mathrm{Q}) = \left|1 - \frac{1}{\mathrm{Q}}\right|, \quad [19]$$

$$\mathrm{E}\left(\mathrm{APE}_{\frac{\bar{s}}{c_4}}\right) = \int_0^\infty \mathrm{f}(\mathrm{Q})\,\mathrm{g}(\mathrm{Q})\,\mathrm{dq}\text{, and} \quad [20]$$

$$\mathrm{SD}\left(\mathrm{APE}_{\frac{\bar{s}}{c_4}}\right) = \left(\mathrm{E}\left(\mathrm{APE}_{\frac{\bar{s}}{c_4}}{}^2\right) - \mathrm{E}^2\left(\mathrm{APE}_{\frac{\bar{s}}{c_4}}\right)\right)^{\frac{1}{2}}. \quad [21]$$

Note that g(Q) in Eq. [20] represents the probability density function for Q. With Eqs. [19]–[21], we can obtain a function that allows us to calculate the required sample size for a given α and Max APE. This function is identical to that in Eq. [14]; however, APE_{S_p} is substituted with $\mathrm{APE}_{\frac{\bar{s}}{c_4}}$.

Results and discussion

In this section, we present the results for the scenarios when practitioners use one baseline sample and multiple baseline samples to estimate the process standard deviation. In each subsection, we first present the expected value of APE and its standard deviation based on different estimators and sample sizes. Those values are calculated based on the formulas in the "Methodology" section and the R codes in the Appendix. Then, we provide some numerical simulations to highlight the between samples variation in APE based on the prescribed sampling plan. We then provide our sample size recommendations based on numerical solutions for the derived functions for n for a given α and Max APE.

Results for the single sampling scenarios

For a single sampling plan, we provide the E(APE) and SD(APE) for both s and $\frac{s}{c_4}$ in Table 2 by using the formula provided in the Appendix. The results demonstrate that as the sample size increases, both E(APE) and SD(APE) decrease. Moreover, smaller E(APE) and SD(APE) values are obtained by using $\frac{s}{c_4}$ instead of s.

Table 2. The expected value and standard deviation of APE for $\hat{\sigma} = s$ and $\hat{\sigma} = s/c_4$.

n	30	40	50	100	150	200	250	300
E(APE_s)	0.1098	0.0935	0.0828	0.0575	0.0466	0.0403	0.0359	0.0328
SD(APE_s)	0.0915	0.0761	0.0664	0.0447	0.0359	0.0309	0.0275	0.0250
E($APE_{\frac{s}{c_4}}$)	0.1084	0.0926	0.0822	0.0572	0.0465	0.0401	0.0358	0.0327
	0.0890	0.0745	0.0653	0.0443	0.0357	0.0307	0.0273	0.0249

In other words, $\frac{s}{c_4}$ can be more efficient than s especially for small sample sizes. It should be noted that the use of the often recommended sample sizes of 30 to estimate the process standard deviation (see Montgomery (2014)) or 50 (see *SEMATECH e-Handbook of Statistical Methods (2012)*) result in E(APE) of at least 8% with a standard deviation that is greater than 6.5%. This means that the estimated values for C_p can vary as much as 25–30% from their true value with these sample sizes, which can result in practitioners drawing wrong conclusions about the capability of their process.

To highlight the variation in APE, consider a situation where a practitioner draws 100 samples of size n to estimate the process standard deviation (either by using s or s/c_4). For the sake of this discussion, let us assume that the practitioner uses $\hat{\sigma} = s$. Since we are focusing on the single sample scenario here, each sample would

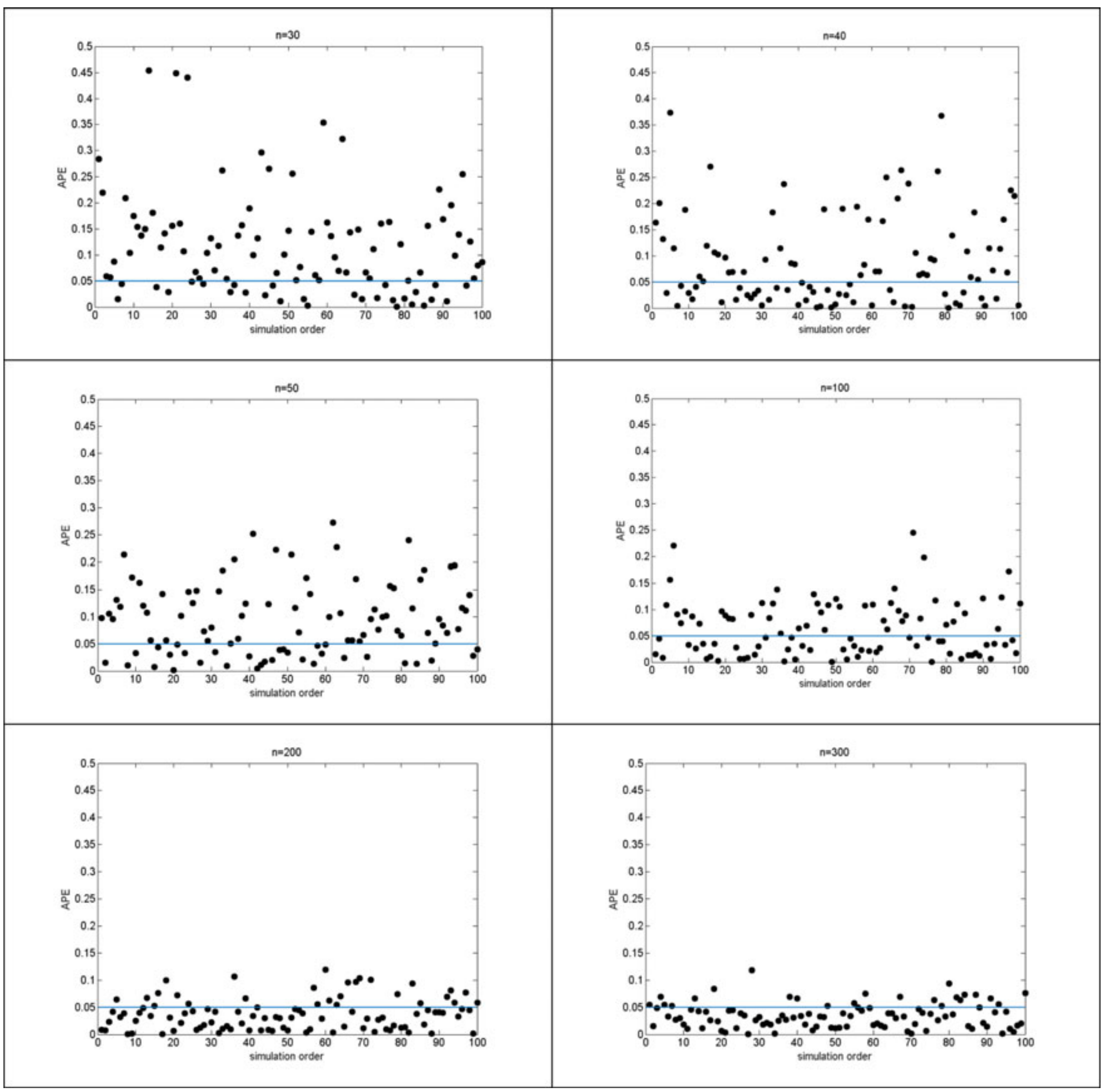

Figure 1. The variation in the APE_s when different values of n are used.

Table 3. Smallest n for $\mathrm{P}(\mathrm{APE}_s < Max\ APE) > 0.95$ and $\mathrm{P}(\mathrm{APE}_{\frac{s}{c_4}} < Max\ APE) > 0.95$.

P($\mathbf{APE}_s < Max\ APE$)	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
MaxAPE$_s$	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
n (smallest sample size)	**4808**	**2140**	**1207**	**774**	**540**	**401**	**306**	**243**	**198**
P($\mathbf{APE}_{\frac{s}{c_4}} < Max\ APE$)	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Max APE$_{\frac{s}{c_4}}$	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
n (smallest sample size)	**4803**	**2137**	**1205**	**773**	**539**	**398**	**305**	**242**	**197**

result in one estimate for C_p. We depict the variation in the APEs associated with this simulation scenario in Figure 1. Note that the purpose of this simulation is to assist readers in visualizing the decrease in APE when n is larger. As expected from Table 2, the variation is reduced as the sample size is increased. Additionally, by setting the value of the Max APE $= 0.05$, we can visualize the function $\mathrm{P}(\mathrm{APE}_s < Max\ APE)$. From the figures, it is clear that the number of practitioners (i.e., points) that are above the Max APE is reduced when n increases. Additionally, it becomes less likely to get larger values for APE. This result is one of the contributions of this article since previous work did not consider the variation in their error metric as a function of n.

Based on Figure 1 and Table 2, we present the smallest sample size needed for different values of Max APE and $\alpha = 0.05$ in Table 3. Note that the sample sizes needed are much larger than what is used in practice if one where to use $\alpha = 0.05$. However, even if different values of α were to be used the required sample size will still be large as shown in Figure 2. Note that the sample size needed varies from 417 to 545 to 774 for $\alpha = 0.15$, $\alpha = 0.10$, and $\alpha = 0.05$, respectively.

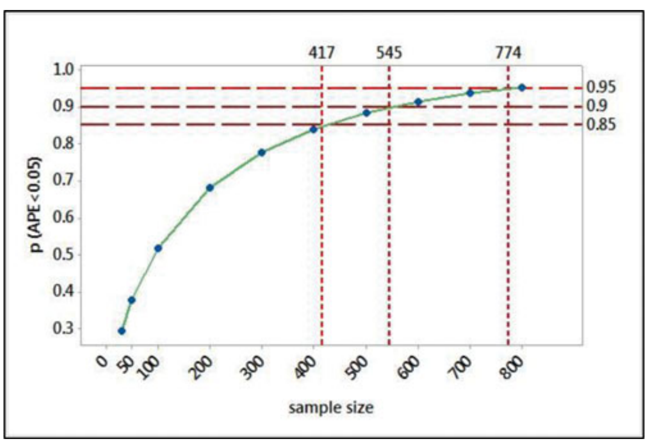

Figure 2. $\mathrm{P}(\mathrm{APE}_s < Max\ APE)$ vs. sample size.

Results for when multiple samples are used

The results for E(APE) and SD(APE) for both S_p and $\bar{s}/c_4$ used (based on different combinations of m and n) to estimate σ are presented in Table 4. Similar to Table 1, the values for E(APE) and SD(APE) decrease as the number of samples (m) and/or the sample size (n) increase. Additionally, the use of S_p is more efficient than $\bar{s}/c_4$, especially when $N = m \times n$ is small, since E(APE) and SD(APE) for S_p are smaller than the corresponding values of $\bar{s}/c_4$. As expected, the values for E(APE) and SD(APE) are dependent on both m and n, i.e., the values are different for the following two scenarios: (a) $n = 5$, $m = 20$ and (b) $n = 10$ and $m = 10$.

To highlight the variation in APE, consider a situation where a practitioner draws 100 samples of size n to estimate the process standard deviation (either by using s_p or $\bar{s}/c_4$). For the sake of this discussion, let us assume that the practitioner uses $\hat{\sigma} = s_p$. Since we are focusing on the multiple sample scenario here, each sample would result in one estimate for C_p.

Similar to Figure 1, we depict the variation in the APEs associated with 100 simulation runs where a practitioner draws m samples of size n in Figure 3. From Table 4, the variation is reduced as m and/or n increases. By setting the value of the Max APE $= 0.05$, we can visualize the function $\mathrm{P}(\mathrm{APE}_{s_p} < Max\ APE)$. From the figures, it is clear that the number of points that are above the Max APE is reduced when m and/or n increases. These results are consistent with Figure 1 and provide insights into why practitioners should consider the variation in the APE.

In Table 5, we provide candidate n and m values when S_p and $\frac{\bar{s}}{c_4}$ are used to estimate σ. For small values of n, the total number of samples required is much smaller when S_p is used. However, m converges for larger values of n. The code provided in the Appendix allows the reader to examine solutions for different values of Max APE and α.

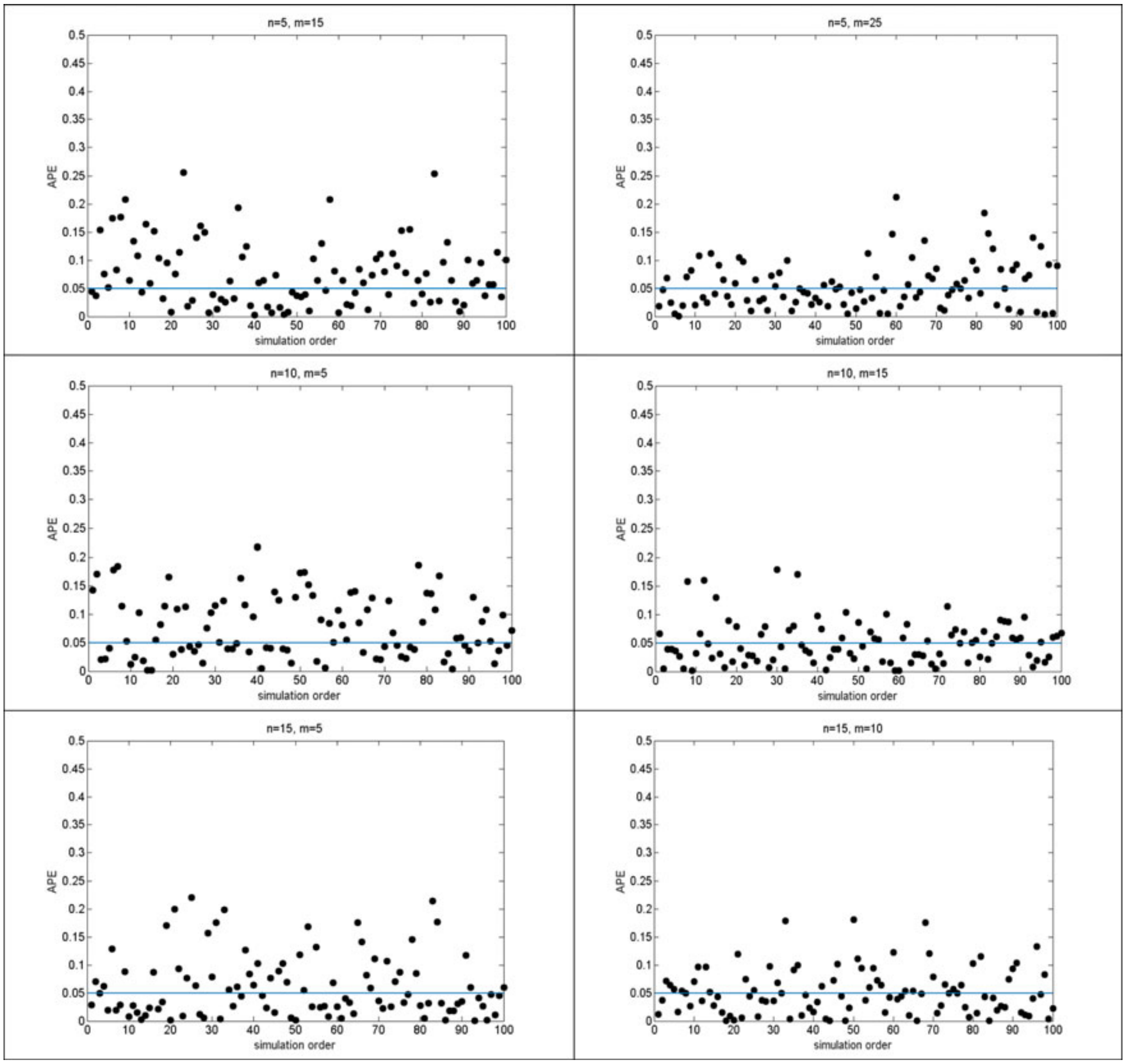

Figure 3. The variation in the APE_{s_p} when different values of *n* and *m* are used.

Table 4. $\mathrm{E(APE)}$ and $\mathrm{SD(APE)}$ for different combination of *n* and *m*.

	N	5	5	5	5	5	5	5	5	10	10	10	10	10
	m	15	20	25	30	40	50	60	75	5	10	15	20	30
S_p	E(APE)	0.0744	0.0641	0.0571	0.0521	0.0450	0.0402	0.0366	0.0327	0.0866	0.0603	0.0490	0.0423	0.0345
	SD(APE)	0.0591	0.0502	0.0445	0.0403	0.0346	0.0308	0.0280	0.0250	0.0698	0.0471	0.0378	0.0325	0.0264
$\frac{\bar{s}}{c_4}$	E(APE)	0.0759	0.0655	0.0585	0.0533	0.0461	0.0411	0.0375	0.0335	0.0869	0.0608	0.0495	0.0428	0.0349
	SD(APE)	0.0599	0.0511	0.0453	0.0411	0.0354	0.0315	0.0287	0.0256	0.0694	0.0473	0.0381	0.0328	0.0266

Table 5. Candidates *m* and *n* values for $\mathrm{P}(\mathrm{APE}_{s_p} < 0.05) > 0.95$ & $\mathrm{P}(\mathrm{APE}_{\frac{\bar{s}}{c_4}} < 0.05) > 0.95$.

P($\mathbf{APE}_{s_p}$ < *Max APE*)	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Max $\mathbf{APE}_{s_p}$	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
n	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**
m	**194**	**86**	**56**	**41**	**33**	**27**	**23**	**20**	**18**
P($\mathbf{APE}_{\frac{\bar{s}}{c_4}}$ < **Max APE**)	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Max $\mathbf{APE}_{\frac{\bar{s}}{c_4}}$	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
n	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**
m	**204**	**88**	**57**	**42**	**33**	**27**	**23**	**20**	**18**

Concluding remarks

In this article, we investigated the effect of estimation error on the process capability index (C_p) using four different estimators of the process standard deviation. We propose using the expected value and standard deviation of the absolute percentage error (APE) to quantify the variation that is seen from one practitioner to another if they were to use a single sample of size n (or m samples of size n) to estimate the process standard deviation. From our calculations, there are four main conclusions one can obtain from this article.

(a) The sample sizes required are generally much larger than the current ones used in industry. For example, the recommendation from the *NIST/SEMATECH e-Handbook of Statistical Methods* is to use $n = 50$. When one sample is used to estimate σ, we recommend n to be in the hundreds based on Table 3. This recommendation would result in increasing the current sample size used in industry by a factor of 10. However, it will minimize the effect of sampling error on C_p.

(b) It is more efficient to use a single sample of size N than m samples of size n (where $N = m \times n$). This can be seen by comparing results from Tables 3 and 5;

(c) If the practitioner has no preference for how to estimate the process standard deviation from a sample of size n, using the estimator s/c_4 is somewhat preferred over s since the standard deviation of the APE is smaller.

(d) The use of APE with the two decision criteria, Max APE and α, provides a simple method to characterize the between samples variation when estimating C_p. Practitioners can easily understand and visualize the impact of their sample size selection on the calculation of C_p by using our tool.

It should be noted that our work assumes that the process output is centered and is normally distributed. These two assumptions are not restrictive. Steiner and Mackay (2005, Ch. 15) provide an algorithm for moving the process center. There are several transformations that can be used to transform non-normal data (e.g., Box-Cox transformation). Perhaps more importantly, these are the assumptions behind using C_p for determining process capability. In this article, our goal is to highlight what sample size is needed (given these assumptions are true) so that the effect of estimation error can be neglected. Future work can address extending our methodology to other process capability indices. It should be noted however that other indices often require the estimation of multiple parameters so the solutions for these should be based on numerical simulations. This might make the development of a tool for practitioners more difficult.

About the authors

Zahra Sedighi Maman is a Ph.D. student in the Department of Industrial and Systems Engineering at Auburn University. Her research interests are in the areas of analytics, data visualization, statistical process control, and quality engineering.

W. Wade Murphy is a Master student in the Department of Industrial and Systems Engineering at Auburn University. His research interests are in: computer programming, data analytics, simulation of complex systems, manufacturing systems design, and operations research.

Saeed Maghsoodloo is a professor emeritus in the Department of Industrial and Systems Engineering at Auburn University. His research interests and publications are in the areas of design of experiments, off-line and on-line quality control, multivariate analysis, regression/correlation and time-series analysis, reliability engineering, and nonparametric statistics. He is a member of ASA and ASQ and is an associate editor of *lIE Transactions* and the *Journal of Manufacturing Systems* and is a reviewer for numerous national and international journals. He is a registered professional engineer in the state of Alabama and is a past member of the editorial board of Quality and Reliability Engineering of lIE Transactions.

Hannane Haji Ahmadi received her MS In Industrial Engineering at Azad Tehran Shomal University, Tehran, Iran. She is currently a research assistant at Iran University of Science and Technology, Tehran, Iran. Her research interests are in reliability engineering, statistical process control, and applied statistics.

Fadel M. Megahed is an assistant professor in the Department of Industrial and Systems Engineering at Auburn University, Auburn, Alabama. His current research focuses on creating new tools to store, organize, analyze, model, and visualize the large heterogeneous data sets associated with modern manufacturing, healthcare and service environments. He is a member of ASQ.

References

Chen, G. 1998. The run length distributions of the R, s and s2 control charts when is estimated. *Canadian Journal of Statistics* 26 (2):311–322.

Chen, C. C., C. M. Lai, and H. Y. Nien. 2010. Measuring process capability index C pm with fuzzy data. *Quality & Quantity* 44 (3):529–535.

Derman, C., and S. Ross. 1995. An improved estimator of σ in quality control. *Probability in the Engineering and Informational Sciences* 9 (03):411–415.

Franklin, L. A. 1999. Sample size determination for lower confidence limits for estimating process capability indices. *Computers & Industrial Engineering* 36 (3):603–614.

Jensen, W. A., L. A. Jones-Farmer, C. W. Champ, and W. H. Woodall. 2006. Effects of parameter estimation on control chart properties: a literature review. *Journal of Quality Technology* 38 (4):349–364.

Jones-Farmer, L. A., W. H. Woodall, S. H. Steiner, and C. W. Champ. 2014. An overview of phase I analysis for process improvement and monitoring. *Journal of Quality Technology* 46 (3):265.

Khattree, R. 1999. On the estimation of σ and the process capability indices Cp and Cpm. *Probability in the Engineering and Informational Sciences* 13 (02):237–250.

Kirmani, S. N. U. A., K. Kocherlakota, and S. Kocherlakota. 1991. Estimation of σ and the process capability index based on subsamples. *Communications in Statistics-Theory and Methods* 20 (1):275–291.

Kotz, S., and N. L. Johnson. 1993. *Process capability indices*. London: CRC Press.

Kumar, M., J. Antony, R. K. Singh, M. K. Tiwari, and D. Perry. 2006. Implementing the Lean Sigma framework in an Indian SME: a case study. *Production Planning and Control* 17 (4):407–423.

Mahmoud, M. A., G. R. Henderson, E. K. Epprecht, and W. H. Woodall. 2010. Estimating the standard deviation in quality-control applications. *Journal of Quality Technology* 42 (4):348.

Montgomery, D. C., G. C. Runger, and N. F. Hubele. 2009. *Engineering statistics*. New York: John Wiley & Sons.

Montgomery, D. C. 2014. *Introduction to statistical quality control*, 7th ed., New York: Wiley & Sons.

NIST/SEMATECH e-Handbook of Statistical Methods. 2012. http://www.itl.nist.gov/div898/handbook/

Patnaik, P. B. 1950. The use of mean range as an estimator of variance in statistical tests. *Biometrika* 37(1/2):78–87.

Pearn, W. L., and M. H. Shu. 2003. Lower confidence bounds with sample size information for Cpm applied to production yield assurance. *International Journal of Production Research* 41 (15):3581–3599.

Steiner, S. H., and R. J. MacKay. 2005. *Statistical engineering: An algorithm for reducing variation in manufacturing processes*, vol. 1. Milwaukee: ASQ Quality Press.

Wu, C. C., and H. L. Kuo. 2004. Sample size determination for the estimate of process capability indices. *International Journal of Information and Management Sciences* 15 (1):1–12.

Zimmer, L. S., N. F. Hubele, and W. J. Zimmer. 2001. Confidence intervals and sample size determination for Cpm. *Quality and Reliability Engineering International* 17 (1):51–68.

Appendix

Finding the sample size based on the different estimators

When s is used to estimate σ

In the section titled "s as an estimator for $\boldsymbol{\sigma}$" we presented Eq. [9] where we stated that an appropriate sample size based on a pre-specified error criterion can be obtained. In this subsection of the Appendix, we start with this equation and show the mathematical manipulations needed to solve for n. Our final equation has n as the only unknown and could be solved using any statistical package. We present our R code for solving this formulation in the next section:

$$P\left(\mathrm{APE}_s < \mathrm{Max\ APE}\right) > 1 - \alpha, \qquad [A1]$$

$$P\left(-\mathrm{Max\ APE} < \frac{Cp - \widehat{Cp}}{Cp} < Max\ APE\right) > 1 - \alpha, \qquad [A2]$$

By substituting C_p and $\widehat{Cp}$ by their corresponding values that are based on Eqs. [1] and [2] and replacing $\hat{\sigma}$ by s, we obtain:

$$P\left(-\mathrm{Max\ APE} < \frac{\frac{USL-LSL}{6\sigma} - \frac{USL-LSL}{6s}}{\frac{USL-LSL}{6\sigma}} < \mathrm{Max\ APE}\right) > 1 - \alpha. \qquad [A3]$$

This will give us:

$$P\left(\frac{1}{1+\mathrm{Max\ APE}} < \frac{s}{\sigma} < \frac{1}{1-\mathrm{Max\ APE}}\right) > 1 - \alpha. \qquad [A4]$$

$$P\left(\frac{1}{(1+\mathrm{Max\ APE})^2} < \frac{s^2}{\sigma^2} < \frac{1}{(1-\mathrm{Max\ APE})^2}\right) > 1 - \alpha. \qquad [A5]$$

By substituting $\frac{s^2}{\sigma^2}$ by its corresponding distribution function, we obtain:

$$P\left(\frac{n-1}{(1+\mathrm{Max\ APE})^2} < \chi^2_{n-1} < \frac{n-1}{(1-\mathrm{Max\ APE})^2}\right) > 1 - \alpha. \qquad [A6]$$

Finally, by determining the α and Max APE, we can solve the equation to find the sample size.

When $\frac{\boldsymbol{s}}{\boldsymbol{c}_4}$ *is used to estimate* $\boldsymbol{\sigma}$

Similar to our discussion for s, lets us consider Eq. [10] in the section titled "s/c4 as an estimator for $\boldsymbol{\sigma}$" to find the appropriate sample size based on a pre-specified error criterion. Then in this subsection of the appendix, we will have:

$$P\left(\mathrm{APE}_{\frac{s}{c_4}} < \text{Max APE}\right) > 1-\alpha\,, \text{ and} \qquad [A7]$$

$$P\left(-\text{Max APE} < \frac{Cp-\widehat{Cp}}{Cp} < \mathit{Max\ APE}\right) > 1-\alpha,$$

By substituting C_p and $\widehat{Cp}$ by their corresponding values, and replacing $\hat{\sigma}$ by $\frac{s}{c_4}$, one could easily obtain:

$$P\left(-\text{Max APE} < \frac{\frac{USL-LSL}{6\sigma}-\frac{USL-LSL}{6\left(\frac{s}{c_4}\right)}}{\frac{USL-LSL}{6\sigma}} < \text{Max APE}\right) > 1-\alpha. \qquad [A8]$$

Then:

$$P\left(\frac{c_4}{1+\text{Max APE}} < \frac{s}{\sigma} < \frac{c_4}{1-\text{Max APE}}\right) > 1-\alpha. \qquad [A9]$$

$$P\left(\left(\frac{c_4}{1+\text{Max APE}}\right)^2 < \frac{s^2}{\sigma^2} < \left(\frac{c_4}{1-\text{Max APE}}\right)^2\right) > 1-\alpha. \qquad [A10]$$

By substituting $\frac{s^2}{\sigma^2}$ by its corresponding probability density function, we obtain:

$$P\left((n-1)\left(\frac{c_4}{1+\text{Max APE}}\right)^2 < \chi^2_{n-1} < (n-1)\times\left(\frac{c_4}{1-\text{Max APE}}\right)^2\right) > 1-\alpha. \qquad [A11]$$

By considering Max APE and a pre-defined α value this final equation has n as the only unknown and could be solved using any statistical package.

When s_p is used to estimate σ

Similar to our procedure for s, consider Eq. [14] in the section titled "Using the pooled sample standard deviation (Sp)" to find the sample size. In this case, we start with this Eq and show the mathematical manipulations needed to solve for n.

$$p\left(\mathrm{APE}_{s_p} < \text{Max APE}\right) > 1-\alpha\,, \text{ and} \qquad [A12]$$

$$P(-\text{Max APE} < \frac{Cp-\widehat{Cp}}{Cp} < \mathit{Max\ APE}) > 1-\alpha,$$

By substituting C_p and $\widehat{Cp}$ by their corresponding values and replacing $\hat{\sigma}$ by s_p, we obtain:

$$P\left(-\text{Max APE} < \frac{\frac{USL-LSL}{6\sigma}-\frac{USL-LSL}{6s_p}}{\frac{USL-LSL}{6\sigma}} < \text{Max APE}\right) > 1-\alpha. \qquad [A13]$$

This will give us:

$$P\left(\frac{1}{1+\text{Max APE}} < \frac{s_p}{\sigma} < \frac{1}{1-\text{Max APE}}\right) > 1-\alpha. \qquad [A14]$$

$$P\left(\frac{1}{(1+\text{Max APE})^2} < \frac{{s_p}^2}{\sigma^2} < \frac{1}{(1-\text{Max APE})^2}\right) > 1-\alpha. \qquad [A15]$$

Then by substituting $\frac{{s_p}^2}{\sigma^2}$ by its corresponding distribution function, we have:

$$P\left(\frac{m(n-1)}{(1+\text{Max APE})^2} < \chi^2_{m(n-1)} < \frac{m(n-1)}{(1-\text{Max APE})^2}\right) > 1-\alpha. \qquad [A16]$$

This final equation has n and m as the only unknown and could be solved using any statistical package.

When $\frac{\bar{s}}{c_4}$ is used to estimate σ

Here, we will follow the same steps as when s is used to estimate σ.

We will use Eq. [A17] to find the appropriate sample size:

$$P\left(\mathrm{APE}_{\frac{\bar{s}}{c_4}} < \text{Max APE}\right) > 1-\alpha\,, \text{ and} \qquad [A17]$$

$$P\left(-\text{Max APE} < \frac{Cp-\widehat{Cp}}{Cp} < \mathit{Max\ APE}\right) > 1-\alpha,$$

By substituting C_p and $\widehat{Cp}$ by their corresponding values and replacing $\hat{\sigma}$ by $\frac{\bar{s}}{c_4}$, we have:

$$P\left(-\text{Max APE} < \frac{\frac{USL-LSL}{6\sigma}-\frac{USL-LSL}{6\left(\frac{\bar{s}}{c_4}\right)}}{\frac{USL-LSL}{6\sigma}} < \mathit{Max\ APE}\right) > 1-\alpha. \qquad [A18]$$

This will give us:

$$P\left(\frac{1}{1+\text{Max APE}} < \frac{\bar{s}}{c_4\sigma} < \frac{1}{1-\text{Max APE}}\right) > 1-\alpha. \qquad [A19]$$

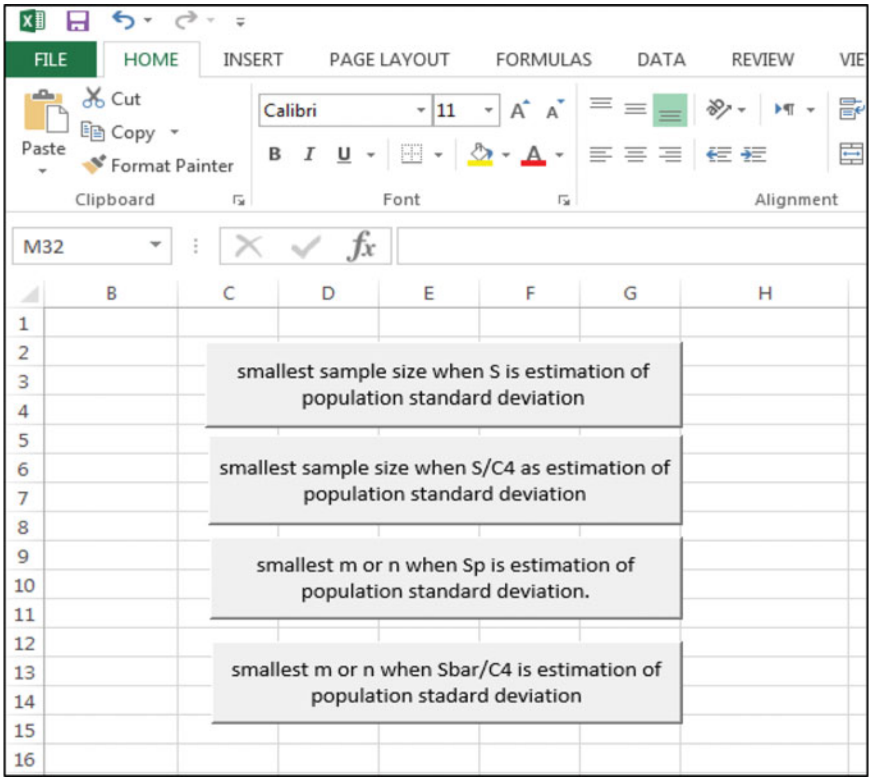

Figure A1. The landing page of the practitioner's toolkit for determining sample size.

By substituting $\frac{\bar{s}}{c_4\sigma}$ by its corresponding distribution function, we obtain:

$$P\left(\frac{1}{1+\text{Max APE}} < \frac{c\chi_\nu}{\sqrt{\nu}} < \frac{1}{1-\text{Max APE}}\right) > 1-\alpha. \quad [A20]$$

This final equation allows us to calculate the required sample n and m for a given α and Max APE.

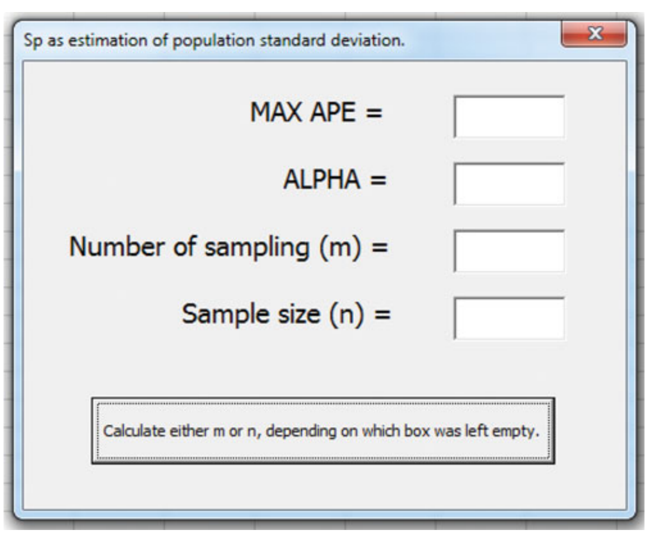

Figure A2. The user-form for the pooled standard deviation.

R codes to find the sample size based on the different estimators

The code used to generate the results discussed in this article uses the R Programming Language (https://www.r-project.org/). To allow researchers to replicate/extend our work, we provide the following link to our code: https://github.com/zahrame/Process-Capability-tool. The reader should note that this shared folder contains four different files; one corresponding to each estimator for σ. To use the code, one should specify the max APE and $1-\alpha$ for the single sample situation. The code will present a suitable sample size based on these constraints. For the multiple samples scenario, the user should also specify either m or n and solve for the other.

An overview of the practitioner's toolkit for sample size determination

Here, we present an excel-based tool that practitioners can use for calculating the sample size (or number of samples) for the different estimators of σ. We provide the Excel based tool at: https://github.com/zahrame/Process-Capability-tool. An overview of the functionality of the tool is provided below to serve as a help document for practitioners.

In the landing page of the tool, we ask the user to select the estimator that they want to use for σ. The four estimators are represented by different buttons as shown in Figure A1. Once the user clicks on any of the four buttons, it will display a particular user-form where he/she can input data for that button and calculate the sample size. For multiple samples, they can also calculate number of samples if they were to provide the sample size as an input (otherwise they should provide the number of samples and solve for the sample size). As an example, we show the user-form generated by selecting the pooled estimator in Figure A2. Based on this information, practitioners can easily determine the sample size needed for their given application. Note that we assume that practitioners have access to Microsoft Excel and that they are familiar with the different estimators for the standard deviation. We believe that these assumptions are reasonable based on our experience with quality practitioners in advanced manufacturing domains.