

This article was downloaded by: [National Chiao Tung University 國立交通大學]

On: 26 April 2014, At: 00:57

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Production Research

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tprs20>

Estimating process yield based on S_{pk} for multiple samples

W. L. Pearn^a & YA Ching Cheng^a

^a Department of Industrial Engineering and Management ,
National Chiao Tung University , Taiwan, Republic of China
Published online: 14 Oct 2010.

To cite this article: W. L. Pearn & YA Ching Cheng (2007) Estimating process yield based on S_{pk} for multiple samples, International Journal of Production Research, 45:1, 49-64, DOI: [10.1080/00207540600600122](https://doi.org/10.1080/00207540600600122)

To link to this article: <http://dx.doi.org/10.1080/00207540600600122>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Estimating process yield based on S_{pk} for multiple samples

W. L. PEARN* and YA CHING CHENG

Department of Industrial Engineering and Management,
National Chiao Tung University, Taiwan, Republic of China

(Revision received January 2006)

Boyles (1994) proposed a process measurement called S_{pk} , which provides an exact measure on the process yield for normal processes. Lee *et al.* (2002) considered an asymptotic distribution for the natural estimator of S_{pk} under a single sample. In this paper, we extend the results for the case of multiple samples. We first compare the yield index S_{pk} with the most commonly used index, C_{pk} , and review some results of S_{pk} under a single sample. Next, we derive the sampling distribution for the estimator \hat{S}'_{pk} of S_{pk} under multiple samples and find that for the same S_{pk} , the variance of \hat{S}'_{pk} would be largest when the process mean is on the centre of specification limits. We calculate the lower bounds for various commonly used quality requirements under the situation with the largest variance of \hat{S}'_{pk} for assurance purposes. To assess the normally approximated distribution of \hat{S}'_{pk} , we simulate with 10 000 replications to generate 10 000 estimates of \hat{S}'_{pk} , calculate their lower bounds, compare with the real (preset) S_{pk} and check the actual type I error. We also compute how many sample sizes are required for the normal approximation to converge to S_{pk} within a designated accuracy. Then, we present a real-world application of the one-cell rechargeable Li-ion battery packs, to illustrate how we apply the lower bounds to actual data collected from factories.

Keywords: Process yield; Multiple samples; Process capability

1. Introduction

Process capability indices (PCIs), which measure the relationship between the manufacturing specifications and the actual process performance, have been widely used in the manufacturing industry providing a numerical measure on whether a process is capable of reproducing items within the specification limits preset

*Corresponding author. Email: wlpearn@mail.nctu.edu.tw

in the factory. Some indices have been explicitly defined as follows:

$$C_a = 1 - \frac{|\mu - m|}{d}, \quad C_p = \frac{USL - LSL}{6\sigma}, \quad C_{pk} = \min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\} = C_a C_p,$$

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}}, \quad C_{pmk} = \min\left\{\frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}}\right\},$$

$$S_{pk} = \frac{1}{3}\Phi^{-1}\left\{\frac{1}{2}\Phi\left(\frac{USL - \mu}{\sigma}\right) + \frac{1}{2}\Phi\left(\frac{\mu - LSL}{\sigma}\right)\right\},$$

where μ is the process mean, σ is the process standard deviation, USL and LSL are the upper and the lower specification limits, respectively, $m = (USL + LSL)/2$, $d = (USL - LSL)/2$, T is the target value, $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution $N(0, 1)$, and $\Phi^{-1}(\cdot)$ is the inverse of $\Phi(\cdot)$.

Numerous capability indices have been proposed to measure the process yield that is an important concern for the manufacturing factories. Process yield is defined as the percentage of processed product units passing the inspection. That is, the product characteristic must fall within the manufacturing tolerance. For processes with two-sided manufacturing specifications, the process yield can be calculated as $\%Yield = F(USL) - F(LSL)$, where $F(\cdot)$ is the cumulative distribution function of the process characteristic. If the process characteristic follows the normal distribution, then the process yield can be alternatively expressed as $\%Yield = \Phi[(USL - \mu)/\sigma] - \Phi[(LSL - \mu)/\sigma]$. Take C_{pk} for example, if $C_{pk} = c$, then the process yield would be in the range of $2\Phi(3c) - 1$ and $\Phi(3c)$, i.e. $2\Phi(3C_{pk}) - 1 \leq \%Yield \leq \Phi(3C_{pk})$ (Boyles 1991). To overcome this shortcoming, Boyles (1994) proposed the yield index, S_{pk} . There is a one-to-one relationship between S_{pk} and the process yield, $\%Yield = 2\Phi(3S_{pk}) - 1$.

Most of the results obtained regarding the statistical properties of estimated capability indices are based on one single sample. However, a common practice in process control is to estimate the process capability indices by using past 'in-control data' from multiple samples, particularly, when a daily-based or weekly-based production control plan is implemented for monitoring process stability. To use estimators based on several small multiple samples and interpret the results as if they were based on a single sample may result in incorrect conclusions. In order to use past in-control data from multiple samples to make decisions regarding process capability, the distribution of the estimated capability index based on multiple samples should be taken into account. When using multiple samples, Kirmani *et al.* (1991) have investigated the distribution of estimators based on the sample standard deviations of the multiple samples. Li *et al.* (1990) have investigated the distribution of estimators of C_p and C_{pk} based on the ranges of the multiple samples. Vännman and Hubele (2003) considered the indices in the class defined by $C_p(u, v)$ and derived the distribution of the estimators of $C_p(u, v)$, when the estimators of the process parameters μ and σ are based on multiple samples.

In this paper, we investigate the behaviour of an estimator of S_{pk} for multiple samples. In the second section, we compare the yield index, S_{pk} , with the most commonly used index, C_{pk} , and review some results of S_{pk} under single sample. In the third section, we derive the sampling distribution for the estimator of S_{pk} under

multiple samples and result in a normal approximation distribution. In the fourth section, we find that the spread of \hat{S}'_{pk} would be largest when the process mean is on the centre of specification limits for the same S_{pk} , so we calculate the lower bounds of S_{pk} from our deriving distribution of \hat{S}'_{pk} based on the situation with the largest variance for conservative. In the fifth section, we show the accuracy of our normal-approximated distribution of \hat{S}'_{pk} by displaying the histograms of lower bounds and the actual type I errors. Finally, we give an application example to describe how to use the lower bounds as listed in our tables.

2. The yield index S_{pk}

We consider a group of five processes as printed in figure 1. For these processes, $USL=36.0$, $LSL=24.0$, and mean $\mu=30.0, 30.5, 31.0, 31.5, 32.0$, standard deviation $\sigma=2.0, 11/6, 5/3, 1.5, 4/3$, respectively (from processes A to E). The C_{pk} value and calculated yield of these five processes are all the same as in table 1(a), and the S_{pk} value and its calculated yield in table 1(b). The ‘Actual Yield’ in table 1(a) and 1(b) is defined by $\Phi((USL - \mu)/\sigma) - \Phi((LSL - \mu)/\sigma)$. We can see that for these five processes the calculated yield of C_{pk} can only guarantee the lower bound yield; however, the calculated yield of S_{pk} value can truly reveal the actual yield of each process.

For single sample, Lee *et al.* (2002) have derived the distribution of an estimator of S_{pk} . The estimator is defined as

$$\hat{S}_{pk} = S_{pk} + \frac{W}{6\sqrt{n}\phi(3S_{pk})} + O_p\left(\frac{1}{n}\right)$$

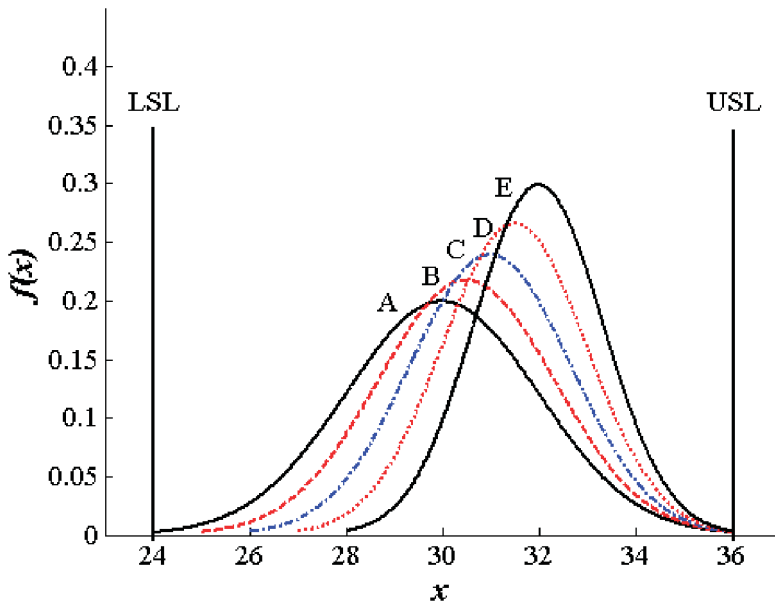


Figure 1. Distribution of five processes with $USL=36.0$, $LSL=24.0$.

Table 1. The C_{pk} value (a), and S_{pk} value (b), calculated yield, and actual yield of five different processes in figure 1.

	μ	σ	C_{pk}	Calculated yield	Actual yield
Process					
A	30.0	2.00	1.0	0.9973	0.9973
B	30.5	1.83	1.0	0.9973	0.9985
C	31.0	1.67	1.0	0.9973	0.9986
D	31.5	1.50	1.0	0.9973	0.9986
E	32.0	1.33	1.0	0.9973	0.9987
	μ	σ	S_{pk}	Calculated yield	Actual yield
Process					
A	30.0	2.00	1.000000	0.9973	0.9973
B	30.5	1.83	1.055311	0.9985	0.9985
C	31.0	1.67	1.067441	0.9986	0.9986
D	31.5	1.50	1.068365	0.9986	0.9986
E	32.0	1.33	1.068385	0.9987	0.9987

where W is normally distributed with a mean of zero and a variance of $a^2 + b^2$,

$$a = \frac{1}{\sqrt{2}} \left\{ \frac{1 - C_{dr}}{C_{dp}} \phi \left(\frac{1 - C_{dr}}{C_{dp}} \right) + \frac{1 + C_{dr}}{C_{dp}} \phi \left(\frac{1 + C_{dr}}{C_{dp}} \right) \right\},$$

$$b = \phi \left(\frac{1 - C_{dr}}{C_{dp}} \right) - \phi \left(\frac{1 + C_{dr}}{C_{dp}} \right),$$

$C_{dr} = (\mu - m)/d$, and $C_{dp} = \sigma/d$. Therefore, \hat{S}_{pk} is asymptotically normal-distributed with mean S_{pk} and variance $(a^2 + b^2)/36n\phi^2(3S_{pk})$. Furthermore, Pearn and Chuang (2004) investigate the accuracy of the natural estimator of S_{pk} , using a simulation technique to find the relative bias and the relative mean square error for some commonly used quality requirement.

Most of the results obtained regarding the statistical properties of estimated capability indices are based on one single sample. However, to use estimators based on several small multiple samples and interpret the results as if they were based on a single sample may result in incorrect conclusions. In order to use past in-control data from multiple samples to make decisions regarding process capability, the distribution of the estimated capability index based on multiple samples should be taken into account. So, in the following we will investigate the sampling distribution of S_{pk} on multiple samples.

3. Estimating S_{pk} under multiple samples

For the case when the studied characteristic of the process is normally distributed and we have m multiple samples where the sample size of the i th sample is n . Let x_{ij} , $i = 1, \dots, m$; $j = 1, \dots, n$, be the characteristic value of the $m \times n$ samples

with mean μ and variance σ^2 . Assume that the process is in statistical control during the time period that the multiple samples are taken. Consider the process is monitored using a \bar{X} -chart together with a S -chart. Then, for each multiple sample, let \bar{x}_i and s_i^2 denote the sample mean and sample variance, respectively, of the i th sample and let N denote the total number of observations, i.e.

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}, \quad s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \quad \text{and} \quad N = \sum_{i=1}^m n = mn.$$

As an estimator of μ , we use the overall sample mean, i.e.

$$\hat{\mu} = \bar{\bar{x}} = \frac{1}{m} \sum_{i=1}^m \bar{x}_i = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n x_{ij}.$$

We consider two ways to compute the variance estimator in estimating S_{pk} (Hubele and Vänman 2004). One estimator of σ^2 is the pooled variance estimator defined as

$$\hat{\sigma}_p^2 = s_p^2 = \frac{1}{mn} \sum_{i=1}^m (n-1)s_i^2 = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2.$$

The other is an un-pooled variance estimator defined as

$$\hat{\sigma}_u^2 = s_u^2 = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{\bar{x}})^2.$$

A natural estimator of S_{pk} is

$$\hat{S}_{pk} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left(\frac{USL - \hat{\mu}}{\hat{\sigma}} \right) + \frac{1}{2} \Phi \left(\frac{\hat{\mu} - LSL}{\hat{\sigma}} \right) \right\}.$$

It is obviously that the sampling distribution of \hat{S}_{pk} is a very complex function of $\hat{\mu}$ and $\hat{\sigma}$. However, a useful approximation could be obtained by the following expansion of S_{pk} . For deriving convenience, we use the notations in Lee's paper:

$$C_{dr} = \frac{\mu - m}{d}, \quad C_{dp} = \frac{\sigma}{d}, \quad \hat{C}_{dr} = \frac{\hat{\mu} - m}{d}, \quad \hat{C}_{dp} = \frac{\hat{\sigma}}{d},$$

and then the estimator of S_{pk} can be rewritten as

$$\hat{S}_{pk} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left(\frac{1 - \hat{C}_{dr}}{\hat{C}_{dp}} \right) + \frac{1}{2} \Phi \left(\frac{1 + \hat{C}_{dr}}{\hat{C}_{dp}} \right) \right\}.$$

Let

$$Z = \sqrt{mn}(\hat{\mu} - \mu) \quad \text{and} \quad Y = \sqrt{mn}(\hat{\sigma}^2 - \sigma^2).$$

We note that $\hat{\mu}$ is a complete sufficient statistic and $\hat{\sigma}^2$ (for either s_p^2 or s_u^2) is an ancillary statistic, so by Basu's theorem Z and Y are independent. Since the first two moments of $\hat{\mu}$ and $\hat{\sigma}^2$ exists, by the Central Limit Theorem, Y converges

to $N(0, 2\sigma^4)$ under both estimators, s_p^2 and s_u^2 , and Z converges to $N(0, \sigma^2)$ as mn goes to infinity. Consequently, by the Taylor's expansion \hat{S}_{pk} can be expressed as

$$\hat{S}_{pk} = S_{pk} + \frac{W}{6\sqrt{mn}\phi(3S_{pk})} + O_p\left(\frac{1}{mn}\right),$$

where

$$W = -\frac{1}{2\sigma^2} Y \left[\frac{1 - C_{dr}}{C_{dp}} \phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) + \frac{1 + C_{dr}}{C_{dp}} \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right] \\ - \frac{1}{\sigma} Z \left[\phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) - \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right]$$

which is normally distributed with mean zero and variance $a^2 + b^2$,

$$a = \frac{1}{\sqrt{2}} \left\{ \frac{1 - C_{dr}}{C_{dp}} \phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) + \frac{1 + C_{dr}}{C_{dp}} \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right\}, \\ b = \phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) - \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right)$$

and ϕ is the pdf of the standard normal distribution (see Appendix I for explicit derivation). We let

$$\hat{S}'_{pk} = \hat{S}_{pk} - O_p\left(\frac{1}{mn}\right)$$

Thus, our \hat{S}'_{pk} is normally distributed, i.e.

$$\hat{S}'_{pk} \sim N\left(S_{pk}, \frac{a^2 + b^2}{36mn\phi^2(3S_{pk})}\right).$$

For testing process performance, we consider the following null and alternative hypotheses:

$$H_0 : S_{pk} \leq c, \quad c \text{ is a specified value. (Process is incapable.)}$$

$$H_1 : S_{pk} > c. \quad \text{(Process is capable.)}$$

The testing statistic is

$$T = \left(\hat{S}_{pk} - c \right) \frac{6\sqrt{mn}\phi(3\hat{S}_{pk})}{\sqrt{\hat{a}^2 + \hat{b}^2}}$$

where \hat{a} and \hat{b} are estimates of a and b , with C_{dr} and C_{dp} replaced by \hat{C}_{dr} and \hat{C}_{dp} , respectively. The null hypothesis H_0 is rejected at α level if $T > z_\alpha$, where z_α is the upper $100\alpha\%$ point of the standard normal distribution. An approximate $1 - \alpha$ confidence interval for S_{pk} is

$$\left(\hat{S}_{pk} - z_{\alpha/2} \frac{\sqrt{\hat{a}^2 + \hat{b}^2}}{6\sqrt{mn}\phi(3\hat{S}_{pk})}, \hat{S}_{pk} + z_{\alpha/2} \frac{\sqrt{\hat{a}^2 + \hat{b}^2}}{6\sqrt{mn}\phi(3\hat{S}_{pk})} \right).$$

4. Lower bounds of S_{pk}

We note that, for the same S_{pk} , the variance of \hat{S}'_{pk} increases as the process mean closes to the centre of the specification limits, and would be largest when the process mean is at the centre of the specification limits. For two processes with the same S_{pk} , i.e. the same process yield, one with process mean away from the centre of the specification limits must have smaller variance in order to have process yield equal to the other. Also, the process with smaller variance would have smaller variance of \hat{S}'_{pk} . Table 2 shows some different processes and corresponding $mnVar(\hat{S}'_{pk})$ with $LSL = 2.0$, $USL = 14.0$, and $S_{pk} = 1.0$.

The following lower bounds displayed in table 3 are calculated under the condition that process mean is on the centre of specification limits for assurance purpose. This approach ensures that the conclusions made based on the lower bounds have the smallest type I error, α , the risk of wrongly concluding an incapable process as capable. When the practitioner wants to know what the least process yield (or say S_{pk}) is, necessary samples could be taken from the 'stable' process to calculate the \hat{S}_{pk} and check the lower bound. The lower bound represents the minimal S_{pk} of the process with $1 - \alpha$ confidence level.

For the convenience of practitioners, we also develop a Matlab program to calculate the lower bounds (see Appendix II). Table 3 shows the lower bounds LB computed from the normal approximation for $\hat{S}_{pk} = 1.0, 1.33, 1.5, 1.67, 2.0$, $n = 5(5)50$, $m = 3, 6$, and $\alpha = 0.05, 0.025, 0.01$. For example, sampling with number of multiple samples $m = 3$ and each of sample size $n = 50$, resulting in sampling estimate $\hat{S}_{pk} = 1.67$, we then conclude that the process has at least $S_{pk} = 1.5221$ with 95% confidence level.

5. Accuracy of the normal approximation

In order to assess the normally approximated distribution of \hat{S}'_{pk} , we simulate with 10 000 replications to generate 10 000 estimates of \hat{S}'_{pk} , calculate their lower bounds, and compare with the real (preset), S_{pk} , for various commonly used quality requirement.

Table 2. Some different processes and corresponding $mnVar(\hat{S}'_{pk})$ with $S_{pk} = 1.0$.

μ	σ	C_{dr}	C_{dp}	a	b	$mnVar(\hat{S}'_{pk})$
8.0000	2.0000	0.0000000	0.3333333	0.0188027	0.0000000	0.499999935
8.0452	1.9995	0.0075315	0.3332483	0.0187932	0.0006001	0.499999793
8.0908	1.9979	0.0151253	0.3329907	0.0187643	0.0012004	0.499995833
8.1371	1.9953	0.0228474	0.3325531	0.0187158	0.0018014	0.499978579
8.1846	1.9915	0.0307724	0.3319220	0.0186472	0.0024033	0.499930778
8.2339	1.9865	0.0389895	0.3310765	0.0185575	0.0030066	0.499826688
8.2857	1.9799	0.0476115	0.3299852	0.0184455	0.0036117	0.499628976
8.3407	1.9716	0.0567897	0.3286013	0.0183095	0.0042190	0.499284463
8.4004	1.9611	0.0667404	0.3268532	0.0181470	0.0048292	0.498717347
8.4668	1.9478	0.0777965	0.3246260	0.0179547	0.0054431	0.497816042
8.5431	1.9303	0.0905222	0.3217203	0.0177273	0.0060618	0.496407586
8.6361	1.9065	0.1060173	0.3177420	0.0174563	0.0066871	0.494199792

Table 3. Approximated LB for various m , \hat{S}_{pk} , $n = 5(5)50$, and $\alpha = 0.05, 0.025, 0.01$.

\hat{S}_{pk}	1.0			1.33			1.5			1.67			2.0			
	m	n		m	n		m	n		m	n		m	n		
3	5	0.7690	0.7364	0.7018	1.0253	0.9819	0.9358	1.1535	1.1046	1.0528	1.2817	1.2274	1.1698	1.5380	1.4729	1.4037
	10	0.8248	0.7980	0.7690	1.0997	1.0640	1.0253	1.2372	1.1970	1.1535	1.3747	1.3301	1.2817	1.6496	1.5961	1.5380
	15	0.8522	0.8287	0.8030	1.1362	1.1050	1.0707	1.2783	1.2431	1.2046	1.4204	1.3813	1.3384	1.7044	1.6575	1.6061
	20	0.8694	0.8482	0.8248	1.1592	1.1309	1.0997	1.3041	1.2723	1.2372	1.4491	1.4137	1.3747	1.7388	1.6964	1.6496
	25	0.8815	0.8620	0.8403	1.1754	1.1493	1.1204	1.3223	1.2930	1.2605	1.4693	1.4367	1.4006	1.7631	1.7240	1.6807
	30	0.8907	0.8725	0.8522	1.1876	1.1633	1.1362	1.3361	1.3088	1.2783	1.4846	1.4542	1.4204	1.7815	1.7450	1.7044
	35	0.8980	0.8808	0.8616	1.1973	1.1744	1.1488	1.3470	1.3212	1.2925	1.4968	1.4681	1.4361	1.7961	1.7617	1.7233
	40	0.9040	0.8876	0.8694	1.2053	1.1835	1.1592	1.3560	1.3315	1.3041	1.5067	1.4795	1.4491	1.8080	1.7753	1.7388
	45	0.9090	0.8934	0.8759	1.2119	1.1912	1.1679	1.3635	1.3401	1.3139	1.5150	1.4890	1.4600	1.8180	1.7868	1.7519
	50	0.9132	0.8983	0.8815	1.2176	1.1977	1.1754	1.3699	1.3475	1.3223	1.5221	1.4972	1.4693	1.8265	1.7966	1.7631
6	5	0.8248	0.7980	0.7690	1.0997	1.0640	1.0253	1.2372	1.1970	1.1535	1.3747	1.3301	1.2817	1.6496	1.5961	1.5380
	10	0.8694	0.8482	0.8248	1.1592	1.1309	1.0997	1.3041	1.2723	1.2372	1.4491	1.4137	1.3747	1.7388	1.6964	1.6496
	15	0.8907	0.8725	0.8522	1.1876	1.1633	1.1362	1.3361	1.3088	1.2783	1.4846	1.4542	1.4204	1.7815	1.7450	1.7044
	20	0.9040	0.8876	0.8694	1.2053	1.1835	1.1592	1.3560	1.3315	1.3041	1.5067	1.4795	1.4491	1.8080	1.7753	1.7388
	25	0.9132	0.8983	0.8815	1.2176	1.1977	1.1754	1.3699	1.3475	1.3223	1.5221	1.4972	1.4693	1.8265	1.7966	1.7631
	30	0.9202	0.9063	0.8907	1.2269	1.2084	1.1876	1.3803	1.3595	1.3361	1.5337	1.5106	1.4846	1.8404	1.8127	1.7815
	35	0.9257	0.9127	0.8980	1.2342	1.2169	1.1973	1.3885	1.3690	1.3470	1.5428	1.5212	1.4967	1.8514	1.8254	1.7961
	40	0.9301	0.9178	0.9040	1.2401	1.2238	1.2053	1.3952	1.3768	1.3560	1.5503	1.5298	1.5067	1.8603	1.8357	1.8080
	45	0.9338	0.9222	0.9089	1.2451	1.2295	1.2119	1.4008	1.3833	1.3634	1.5565	1.5370	1.5150	1.8677	1.8444	1.8179
	50	0.9370	0.9259	0.9132	1.2493	1.2345	1.2176	1.4056	1.3888	1.3698	1.5618	1.5432	1.5221	1.8741	1.8518	1.8265

Figure 2(a)–(d) shows histograms of lower bounds each of 10 000 replications with $\alpha=0.05$, $m=3$, $n=50$, $S_{pk}=1.00$, 1.33, 1.50, and 1.67, respectively. Table 4 displays the actual type I errors for various m , n , S_{pk} , and each with 10 000 simulated lower bounds.

The results in table 4 show that when $m=12$, $n=50$, the confidence level of the normal approximation is almost equal to the preset $1-\alpha$ (the confidence levels are all greater than 94%). As we know, the simulation results are in large variation, and by Central Limit Theorem the average is in small variation, so we calculate the average of the lower bounds and compare to the real S_{pk} . Table 5 shows the ratios of the average lower bounds relative to the real S_{pk} . It is noted that no matter what the real S_{pk} is, the ratios of \overline{LB}/S_{pk} are almost equal with the same m and n . Thus, it is reasonable to estimate the true S_{pk} from the ratios. For example, when $m=3$ and $n=200$, practitioners can repeat the sampling procedure, obtain the average lower bound, and estimate the real S_{pk} by $\overline{LB}/0.9558$.

We further consider how many sample size n should be taken to ensure that the sampling estimator is closed enough to the real S_{pk} within a designated accuracy ε (Pearn *et al.* 2004b). Table 6 displays the sample sizes required for the normal approximation to converge to the real S_{pk} within a designated accuracy ε less than 0.10, 0.09, 0.08, 0.07, 0.06, 0.05, 0.04, 0.03, 0.02, 0.01, respectively,

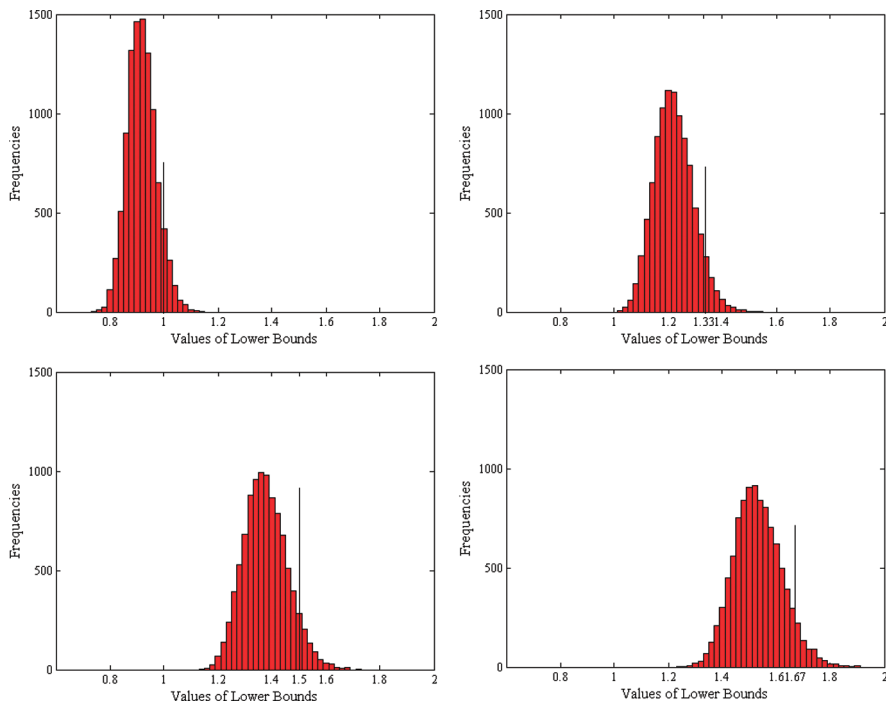


Figure 2. Histogram of 10 000 lower bounds with (a) $S_{pk} = 1.00$; (b) $S_{pk} = 1.33$; (c) $S_{pk} = 1.50$; (d) $S_{pk} = 1.67$.

Table 4. Simulated type I errors α for various m , n , and S_{pk} with 10 000 lower bounds.

m	S_{pk}	$n = 10$	$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 150$	$n = 200$
1	1.00	0.1520	0.1156	0.1015	0.0812	0.0788	0.0680	0.0695
	1.33	0.1593	0.1147	0.0987	0.0839	0.0750	0.0680	0.0694
	1.50	0.1634	0.1161	0.1015	0.0869	0.0707	0.0677	0.0656
	1.67	0.1629	0.1132	0.1053	0.0842	0.0764	0.0673	0.0663
	2.00	0.1691	0.1188	0.1012	0.0845	0.0737	0.0701	0.0684
2	1.00	0.1126	0.0897	0.0817	0.0748	0.0655	0.0628	0.0602
	1.33	0.1124	0.0901	0.0823	0.0715	0.0677	0.0632	0.0614
	1.50	0.1148	0.0940	0.0842	0.0706	0.0684	0.0692	0.0663
	1.67	0.1158	0.0939	0.0823	0.0754	0.0702	0.0588	0.0584
	2.00	0.1223	0.0904	0.0844	0.0715	0.0639	0.0613	0.0641
3	1.00	0.1017	0.0838	0.0742	0.0692	0.0629	0.0579	0.0563
	1.33	0.1011	0.0825	0.0726	0.0651	0.0671	0.0560	0.0556
	1.50	0.1016	0.0791	0.0784	0.0638	0.0632	0.0600	0.0567
	1.67	0.1048	0.0854	0.0681	0.0681	0.0624	0.0586	0.0556
	2.00	0.1063	0.0805	0.0772	0.0693	0.0664	0.0644	0.0576
6	1.00	0.0816	0.0693	0.0705	0.0634	0.0619	0.0573	0.0555
	1.33	0.0831	0.0750	0.0669	0.0576	0.0624	0.0545	0.0578
	1.50	0.0832	0.0727	0.0677	0.0634	0.0599	0.0611	0.0555
	1.67	0.0861	0.0744	0.0673	0.0576	0.0609	0.0603	0.0575
	2.00	0.0763	0.0741	0.0699	0.0681	0.0626	0.0568	0.0558
9	1.00	0.0739	0.0675	0.0687	0.0573	0.0583	0.0561	0.0553
	1.33	0.0752	0.0655	0.0652	0.0546	0.0590	0.0546	0.0550
	1.50	0.0762	0.0671	0.0610	0.0628	0.0528	0.0561	0.0535
	1.67	0.0804	0.0688	0.0616	0.0623	0.0597	0.0577	0.0554
	2.00	0.0748	0.0713	0.0641	0.0636	0.0609	0.0533	0.0498
12	1.00	0.0707	0.0643	0.0593	0.0569	0.0553	0.0534	0.0557
	1.33	0.0751	0.0657	0.0559	0.0570	0.0625	0.0568	0.0521
	1.50	0.0671	0.0652	0.0641	0.0590	0.0559	0.0559	0.0545
	1.67	0.0728	0.0682	0.0625	0.0597	0.0587	0.0577	0.0554
	2.00	0.0705	0.0599	0.0645	0.0575	0.0542	0.0538	0.0520

and the derivation is briefly done as follows:

$$\begin{aligned}
 \Pr\left\{\left|\hat{S}'_{pk} - S_{pk}\right| \leq \varepsilon\right\} &\geq 1 - \alpha \Rightarrow \Pr\left\{\frac{\hat{S}'_{pk} - S_{pk}}{\sqrt{\text{Var}(\hat{S}'_{pk})}} \leq \frac{\varepsilon}{\sqrt{\text{Var}(\hat{S}'_{pk})}}\right\} \geq 1 - \frac{\alpha}{2} \\
 &\Rightarrow \frac{\varepsilon}{\sqrt{\text{Var}(\hat{S}'_{pk})}} \geq \Phi^{-1}(1 - \alpha/2) \\
 &\Rightarrow \frac{a^2 + b^2}{36mn\phi^2(3S_{pk})} \leq \frac{\varepsilon^2}{[\Phi^{-1}(1 - \alpha/2)]^2} \\
 &\Rightarrow mn \geq \frac{(a^2 + b^2)[\Phi^{-1}(1 - \alpha/2)]^2}{36\phi^2(3S_{pk})\varepsilon^2},
 \end{aligned}$$

Table 5. Ratios of the average of 10 000 lower bounds and the real S_{pk} , i.e. \overline{LB}/S_{pk} .

m	S_{pk}	$n = 10$	$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 150$	$n = 200$
1	1.00	0.8056	0.8286	0.8483	0.8726	0.9030	0.9178	0.9275
	1.33	0.8122	0.8287	0.8487	0.8729	0.9032	0.9179	0.9273
	1.50	0.8156	0.8302	0.8505	0.8730	0.9029	0.9178	0.9271
	1.67	0.8171	0.8312	0.8511	0.8732	0.9032	0.9186	0.9274
	2.00	0.8239	0.8341	0.8510	0.8739	0.9038	0.9177	0.9278
2	1.00	0.8283	0.8611	0.8810	0.9030	0.9274	0.9388	0.9460
	1.33	0.8314	0.8636	0.8802	0.9025	0.9276	0.9396	0.9472
	1.50	0.8319	0.8627	0.8817	0.9021	0.9285	0.9400	0.9469
	1.67	0.8304	0.8628	0.8815	0.9037	0.9277	0.9391	0.9466
	2.00	0.8358	0.8629	0.8816	0.9020	0.9277	0.9394	0.9471
3	1.00	0.8496	0.8820	0.8982	0.9179	0.9390	0.9493	0.9552
	1.33	0.8493	0.8806	0.8979	0.9170	0.9402	0.9497	0.9555
	1.50	0.8484	0.8810	0.8983	0.9179	0.9390	0.9494	0.9555
	1.67	0.8506	0.8815	0.8976	0.9176	0.9393	0.9493	0.9558
	2.00	0.8516	0.8812	0.8987	0.9185	0.9401	0.9498	0.9556
6	1.00	0.8807	0.9095	0.9241	0.9394	0.9563	0.9632	0.9682
	1.33	0.8803	0.9100	0.9239	0.9388	0.9559	0.9634	0.9684
	1.50	0.8810	0.9106	0.9241	0.9394	0.9561	0.9636	0.9683
	1.67	0.8825	0.9096	0.9241	0.9387	0.9556	0.9635	0.9685
	2.00	0.8819	0.9102	0.9240	0.9400	0.9563	0.9635	0.9679
9	1.00	0.8988	0.9245	0.9369	0.9499	0.9630	0.9697	0.9734
	1.33	0.8994	0.9236	0.9357	0.9495	0.9637	0.9700	0.9735
	1.50	0.8988	0.9248	0.9361	0.9499	0.9634	0.9699	0.9736
	1.67	0.8995	0.9244	0.9363	0.9500	0.9638	0.9702	0.9737
	2.00	0.8996	0.9247	0.9369	0.9499	0.9635	0.9697	0.9735
12	1.00	0.9104	0.9329	0.9442	0.9561	0.9681	0.9736	0.9770
	1.33	0.9100	0.9330	0.9434	0.9558	0.9683	0.9736	0.9769
	1.50	0.9093	0.9333	0.9441	0.9557	0.9682	0.9738	0.9771
	1.67	0.9095	0.9337	0.9441	0.9560	0.9682	0.9736	0.9773
	2.00	0.9106	0.9323	0.9446	0.9556	0.9677	0.9735	0.9772

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution $N(0, 1)$, $\Phi^{-1}(\cdot)$ is the inverse of $\Phi(\cdot)$, and $\phi(\cdot)$ is the probability density function of $N(0, 1)$.

For example, for $m=9$, $S_{pk}=1.33$ with risk $\alpha=0.05$, a sample size of $n \geq 3795$ ensures that the difference between the sampling \hat{S}_{pk} and the real S_{pk} is smaller than 0.01. Thus, if the sampling $\hat{S}_{pk}=1.33$, then we can conclude that the actual performance $S_{pk} > 1.32$ with 95% confidence level. This convergence investigated is not for practical purpose, but to illustrate the behaviour and the rate of convergence for the normal approximation.

6. An application example

The integrated circuits (IC) industry has been the most popular industry of previous years. Products of integrated circuits are various types such as office automation equipment (copiers, facsimile machines, printers, etc.), vending machines, banking terminals, CD or DVD players, battery chargers, etc. We investigated a company

Table 6. Sample sizes required for the normal approximation to converge with $\alpha = 0.05$.

m	S_{pk}	Designated accuracy, ε									
		0.10	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01
1	1.00	193	238	301	392	534	769	1201	2135	4802	19208
	1.33	342	422	534	697	949	1366	2135	3795	8537	34147
	1.50	433	534	676	882	1201	1729	2702	4802	10805	43217
	1.67	534	659	834	1089	1483	2135	3335	5929	13339	53354
	2.00	769	949	1201	1568	2135	3074	4802	8537	19208	76830
2	1.00	97	119	151	196	267	385	601	1068	2401	9604
	1.33	171	211	267	349	475	683	1068	1898	4269	17074
	1.50	217	267	338	441	601	865	1351	2401	5403	21609
	1.67	267	330	417	545	742	1068	1668	2965	6670	26677
	2.00	385	475	601	784	1068	1537	2401	4269	9604	38415
3	1.00	65	80	101	131	178	257	401	712	1601	6403
	1.33	114	141	178	233	317	456	712	1265	2846	11383
	1.50	145	178	226	294	401	577	901	1601	3602	14406
	1.67	178	220	278	363	495	712	1112	1977	4447	17785
	2.00	257	317	401	523	712	1025	1601	2846	6403	25610
6	1.00	33	40	51	66	89	129	201	356	801	3202
	1.33	57	71	89	117	159	228	356	633	1423	5692
	1.50	73	89	113	147	201	289	451	801	1801	7203
	1.67	89	110	139	182	248	356	556	989	2224	8893
	2.00	129	159	201	262	356	513	801	1423	3202	12805
9	1.00	22	27	34	44	60	86	134	238	534	2135
	1.33	38	47	60	78	106	152	238	422	949	3795
	1.50	49	60	76	98	134	193	301	534	1201	4802
	1.67	60	74	93	121	165	238	371	659	1483	5929
	2.00	86	106	134	175	238	342	534	949	2135	8537
12	1.00	17	20	26	33	45	65	101	178	401	1601
	1.33	29	36	45	59	80	114	178	317	712	2846
	1.50	37	45	57	74	101	145	226	401	901	3602
	1.67	45	55	70	91	124	178	278	495	1112	4447
	2.00	65	80	101	131	178	257	401	712	1601	6403

in Taiwan manufacturing one-cell rechargeable Li-ion battery packs which have advantages of low current consumption, high withstand voltage, high accuracy voltage detection, over current and short circuit protection, and wide operating temperature range. Among the advanced features, the most important one is the high accuracy voltage detection. Once the voltage detector falls down, the lifetime or reliability of the Li-ion battery pack will be discounted. The preset upper and lower specification limits of the over charge detector are $USL = 4.40$ V, $LSL = 4.30$ V, and target value is set to $T = 4.35$ V.

The electrical characteristic data of 12 multiple samples each of size 50 are collected. The sample mean, \bar{x}_i , and sample standard deviation, s_i , for the 12 samples are listed in table 7. To make the estimate of \hat{S}_{pk} meaningful, we first check if the characteristic data collected from the process is in control and normally distributed. For those 12 samples of size 50 each, the Kolmogorov–Smirnov test confirms the sample normal with a P value > 0.15 . That is, it is reasonable to assume that the data collected from the process is normally distributed. Then, we construct the $\bar{X} - S$ charts to check if the process is under statistical control. Figure 3 shows the $\bar{X} - S$

Table 7. The collected electrical characteristic data of 12 samples each of size 50.

Sub-sample	1	2	3	4	5	6	7	8	9	10	11	12
\bar{x}_i	4.3526	4.3483	4.3544	4.3490	4.3563	4.3542	4.3482	4.3537	4.3535	4.3505	4.3476	4.3502
s_i	0.0133	0.0120	0.0124	0.0093	0.0104	0.0114	0.0119	0.0174	0.0126	0.0112	0.0104	0.0102

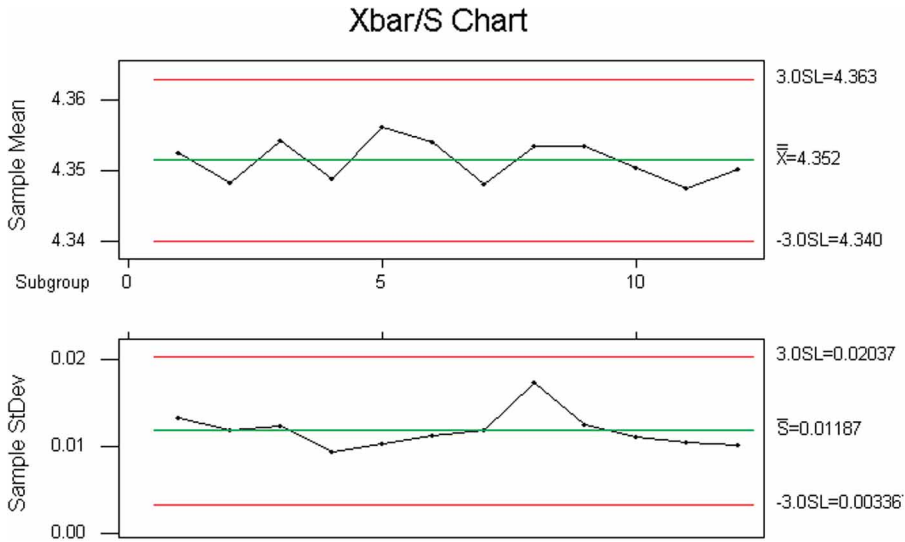


Figure 3. The $\bar{X}-S$ charts based on the collected 12 samples each of size 50.

charts based on the collected samples. Since all sample points are within the control limits without any special pattern, we can conclude that the process is in control. Therefore, we consider the process stable and then proceed to the capability measurements.

Suppose the minimal precision requirement for this process is set to $S_{pk} = 1.0$. We calculate the overall sample mean $\bar{\bar{x}}_i = 4.35154$, the pooled sample standard deviation $s_p = \sqrt{s_p^2} = 0.01192$, the un-pooled sample standard deviation $s_u = \sqrt{s_u^2} = 0.01225$, and

$$\begin{aligned} \hat{S}_{pk} &= \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left(\frac{USL - \hat{\mu}}{\hat{\sigma}} \right) + \frac{1}{2} \Phi \left(\frac{\hat{\mu} - LSL}{\hat{\sigma}} \right) \right\} \\ &= \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left(\frac{4.40 - 4.35154}{0.01192} \right) + \frac{1}{2} \Phi \left(\frac{4.35154 - 4.30}{0.01192} \right) \right\} \end{aligned}$$

or

$$\begin{aligned} &\frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi \left(\frac{4.40 - 4.35154}{0.01225} \right) + \frac{1}{2} \Phi \left(\frac{4.35154 - 4.30}{0.01225} \right) \right\} \\ &= 1.3871 \text{ or } 1.3503. \end{aligned}$$

We run the program in Appendix II to find the lower bound as 1.3242 (or 1.2890). Thus, we conclude that the true value of the process capability S_{pk} would be no less than 1.3242 (or 1.2890) with 95% confidence level.

To estimate the real S_{pk} , the factory manager could implement a weekly based control system by repeating the sampling procedure for consecutive, say 10 weeks, then calculate the average lower bounds. For example, the 10 weeks lower bounds result $\overline{LB} = 1.4527$. Refer to table 5, the biggest ratio of \overline{LB}/S_{pk} for $m = 12$ and $n = 50$ is 0.9561, then he can estimate the real $S_{pk} = 1.4527/0.9561 \doteq 1.52$. The corresponding process yield then could be estimated as 0.999994885, or equally, fraction of defectives is 5.115 ppm.

7. Conclusion

How to measure process performance in the manufacturing industries is a major concern for the factory managers, and process yield is the most common and standard criteria. The capability index, S_{pk} , provides an exact yield measure (rather than yield range) for normal processes. Many researches supporting the use of capability indices have focused on processes with a single large representative sample. However, in practice the process performance is monitored by collecting multiple samples periodically. Our paper considered this type of realistic data structure and investigated the sampling distribution of \hat{S}'_{pk} based on multiple samples.

We note that for processes with the same specification limits and process yield, the variance of \hat{S}'_{pk} would be largest while process mean is on the centre of the specification limits, so we compute the lower bounds of S_{pk} under such condition for assurance purpose. It is noted that the lower bounds calculated by normal approximation distribution have larger risk of α especially for smaller total sample size $m \times n$. As mentioned before, sampling with number of multiple samples $m = 12$, and number of sample size $n = 50$ is suggested to use the lower bounds from normal approximation to have the almost $1 - \alpha$ confidence level. Furthermore, if the sampling replications are allowed, practitioners can even estimate the real process capability from the average of lower bounds dividing by the ratio (i.e. $S_{pk} \doteq \overline{LB}/\text{ratio}$).

Appendix

Appendix I. Taylor's expansion of \hat{S}_{pk}

Before our derivation, we need to define some notations:

$$C_{dr} = \frac{\mu - m}{d}, \quad C_{dp} = \frac{\sigma}{d}, \quad \hat{C}_{dr} = \frac{\hat{\mu} - m}{d}, \quad \hat{C}_{dp} = \frac{\hat{\sigma}}{d},$$

$$Z = \sqrt{mn}(\hat{\mu} - \mu) \quad \text{and} \quad Y = \sqrt{mn}(\hat{\sigma}^2 - \sigma^2)$$

By the Central Limit Theorem, Y converges to $N(0, 2\sigma^4)$ under both estimators, s_p^2 and s_u^2 , and Z converges to $N(0, \sigma^2)$ as mn goes to infinity.

$$\hat{C}_{dr} = C_{dr} + \frac{Z}{\sigma\sqrt{mn}}(C_{dp}), \quad 1 - \hat{C}_{dr} = (1 - C_{dr}) - \frac{Z}{\sigma\sqrt{mn}}(C_{dp}),$$

$$1 + \hat{C}_{dr} = (1 + C_{dr}) + \frac{Z}{\sigma\sqrt{mn}}(C_{dp}),$$

$$\hat{C}_{dp} = C_{dp}\sqrt{1 + \frac{Y}{\sigma^2\sqrt{mn}}}, \quad \frac{1}{\hat{C}_{dp}} = \frac{1}{C_{dp}} \left[1 - \frac{Y}{2\sigma^2\sqrt{mn}} + O_p\left(\frac{1}{mn}\right) \right],$$

$$\frac{1 - \hat{C}_{dr}}{\hat{C}_{dp}} = \frac{1 - C_{dr}}{C_{dp}} + \frac{1}{\sqrt{mn}} \left[\frac{-Z}{\sigma} - \left(\frac{1 - C_{dr}}{C_{dp}} \right) \frac{Y}{2\sigma^2} \right] + O_p\left(\frac{1}{mn}\right),$$

$$\frac{1 + \hat{C}_{dr}}{\hat{C}_{dp}} = \frac{1 + C_{dr}}{C_{dp}} + \frac{1}{\sqrt{mn}} \left[\frac{Z}{\sigma} - \left(\frac{1 + C_{dr}}{C_{dp}} \right) \frac{Y}{2\sigma^2} \right] + O_p\left(\frac{1}{mn}\right),$$

$$\Phi\left(\frac{1 - \hat{C}_{dr}}{\hat{C}_{dp}}\right) = \Phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) + \frac{1}{\sqrt{mn}} \phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) \left[\frac{-Z}{\sigma} - \left(\frac{1 - C_{dr}}{C_{dp}} \right) \frac{Y}{2\sigma^2} \right] + O_p\left(\frac{1}{mn}\right)$$

$$\Phi\left(\frac{1 + \hat{C}_{dr}}{\hat{C}_{dp}}\right) = \Phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) + \frac{1}{\sqrt{mn}} \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \left[\frac{Z}{\sigma} - \left(\frac{1 + C_{dr}}{C_{dp}} \right) \frac{Y}{2\sigma^2} \right] + O_p\left(\frac{1}{mn}\right)$$

$$\Phi\left(\frac{1 - \hat{C}_{dr}}{\hat{C}_{dp}}\right) + \Phi\left(\frac{1 + \hat{C}_{dr}}{\hat{C}_{dp}}\right) = \Phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) + \Phi\left(\frac{1 + C_{dr}}{C_{dp}}\right)$$

$$+ \frac{1}{\sqrt{mn}} \frac{-Z}{\sigma} \left[\phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) - \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right]$$

$$+ \frac{1}{\sqrt{mn}} \frac{-Y}{2\sigma^2} \left[\left(\frac{1 - C_{dr}}{C_{dp}} \right) \phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) + \left(\frac{1 + C_{dr}}{C_{dp}} \right) \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right]$$

$$+ O_p\left(\frac{1}{mn}\right)$$

Since the estimator of S_{pk} can be rewritten as

$$\hat{S}_{pk} = \frac{1}{3} \Phi^{-1} \left\{ \frac{1}{2} \Phi\left(\frac{1 - \hat{C}_{dr}}{\hat{C}_{dp}}\right) + \frac{1}{2} \Phi\left(\frac{1 + \hat{C}_{dr}}{\hat{C}_{dp}}\right) \right\},$$

applying the Taylor's expansion

$$\Phi^{-1}(x + y) = \Phi^{-1}(x) + \frac{y}{\phi(\Phi^{-1}(x))} + O_p(y^2),$$

we can obtain

$$3\hat{S}_{pk} = 3S_{pk} + \frac{W}{2\sqrt{mn}\phi(3S_{pk})} + O_p\left(\frac{1}{mn}\right),$$

where

$$W = -\frac{Y}{2\sigma^2} \left[\frac{1 - C_{dr}}{C_{dp}} \phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) + \frac{1 + C_{dr}}{C_{dp}} \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right] - \frac{Z}{\sigma} \left[\phi\left(\frac{1 - C_{dr}}{C_{dp}}\right) - \phi\left(\frac{1 + C_{dr}}{C_{dp}}\right) \right].$$

Thus

$$\hat{S}_{pk} = S_{pk} + \frac{W}{6\sqrt{mn}\phi(3S_{pk})} + O_p\left(\frac{1}{mn}\right).$$

Appendix II. Calculation of lower bounds

```

alpha=0.05;
m=12; % number of sub-samples
n=50; % number of sample size
S=1.0; % value of Spk head
Spk=S;
for i=1:1:100000
Spk=Spk-0.0001;
Cdp=1/(3*Spk);
a=sqrt(2)*3*Spk*normpdf(3*Spk);
p=normcdf((S-Spk)/sqrt(a*a/(36*m*n*normpdf(3*Spk)*normpdf(3*Spk))));
if p>(1-alpha) break;
end; end
fprintf('The true value of the process capability Spk is no less
than%g', Spk)

```

References

- Boyles, R.A., The Taguchi capability index. *J. Qual. Technol.*, 1991, **23**, 17–26.
- Boyles, R.A., Process capability with asymmetric tolerances. *Commun. Stat. Simulat.*, 1994, **23**(3), 615–643.
- Hubele, N.F. and Vännman, K., The effect of pooled and un-pooled variance estimators on C_{pm} when using subsamples. *J. Qual. Tech.*, 2004, **36**(2), 207–222.
- Lee, J.C., Hung, H.N., Pearn, W.L. and Kueng, T.L., On the distribution of the estimated process yield index S_{pk} . *Quality and Reliability Engineering International*, 2002, **18**(2), 111–116.
- Li, H., Owen, D.B. and Borrego, S.A., Lower confidence limits on process capability indices based on the range. *Commun. Stat. Simulat.*, 1990, **19**(1), 1–24.
- Kirmani, S.N.U.A., Kocherlakota, K. and Kocherlakota, S., Estimation of σ and the process capability index based on subsamples. *Commun. Stat. Theory*, 1991, **20**, 275–291.
- Pearn, W.L. and Chuang, C.C., Accuracy analysis of the estimated process yield based on S_{pk} . *Qual. Reliab. Eng. Int.*, 2004a, **38**(1), 95–111.
- Pearn, W.L., Lin, G.H. and Wang, K.H., Normal approximation to the distribution of the estimated yield index S_{pk} . *Qual. Quant.*, 2004b, **38**(1), 95–111.
- Vännman, K. and Hubele, N.F., Distributional properties of estimated capability indices based on subsamples. *Qual. Reliab. Eng. Int.*, 2003, **19**, 111–128.