Yugoslav Journal of Operations Research
Volume 19 (2009), Number 2, 215-223
DOI:10.2298/YUJOR0902215L

A RESEARCH NOTE ON THE ESTIMATED INCAPABILITY INDEX

Gu-Hong LIN

Department of Industrial Engineering and Management, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan, ROC
e-mail: ghlin@cc.kuas.edu.tw

Received: March 2008 / Accepted: Februaruy 2009

Abstract: Process incapability index, which provides an uncontaminated separation between information concerning the process accuracy and the process precision, has been proposed to the manufacturing industry for measuring process performance. Contributions concerning the estimated incapability index have focused on single normal process in existing quality and statistical literature. However, the contaminated model is more appropriate for real-world cases with multiple manufacturing processes where the raw material, or the equipment may not be identical for each manufacturing process. Investigations based on contaminated normal processes are considered. Sampling distributions and *r*-th moments of the estimated index are derived. The proposed model will facilitate quality engineers on process monitor and performance assessment.

Keywords: Contaminated normal process; incapability index

1. INTRODUCTION

Process capability indices, whose purpose is to provide numerical measures on whether or not a manufacturing process is capable of reproducing items satisfying the quality requirements preset by the customers, the product designers, have received substantial research attention in the quality control and statistical literature. The three basic capability indices C_p, C_a and C_{pk}, have been defined as (e.g. Kane, 1986; Pearn et al., 1998; Lin, 2006a):

$$C_a = 1 - \frac{|\mu - m|}{d}, \tag{1}$$

$$C_p = \frac{USL - LSL}{6\sigma}, \tag{2}$$

$$C_{pk} = \min\left\{\frac{USL-\mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\} \tag{3}$$

where USL and LSL are the upper and lower specification limits preset by the customers, the product designers, μ is the process mean, σ is the process standard deviation, $m = (USL + LSL)/2$ and $d = (USL + LSL)/2$ are the mid-point and half length of the specification interval, respectively.

The index C_p reflects only the magnitude of the process variation relative to the specification tolerance and, therefore, is used to measure process potential. The index C_a measures the degree of process centering (the ability to cluster around the center) and is referred as the process accuracy index. The index C_{pk} takes into account process variation as well as the location of the process mean. The natural estimators of C_p, C_a, and C_{pk} can be obtained by substituting the sample mean $\bar{X} = \sum_{i=1}^{n} X_i / n$ for μ and the sample variance $S_{n-1}^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 /(n-1)$ for C_p in the expressions (1), (2), and (3). Chou et al. (1989), Kotz et al. (1993), Pearn et al. (1998), and Lin (2006a) investigated the statistical properties and the sampling distributions of the natural estimators of C_p, C_a, and C_{pk}.

Boyles (1991) noted that C_{pk} is a yield-based index. In fact, the design of C_{pk} is independent of the target value T and C_{pk} can fail to account for process targeting (the ability to cluster around the target). For this reason, Chan et al. (1988) developed the index C_{pm} to take the process targeting issue into consideration. The index C_{pm} is defined as the following:

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

Johnson (1992) pointed out that the index C_{pm} is not originally designed to provide an exact measure on the number of non-conforming items, but a loss-based index.

For processes with asymmetric tolerance $(T \neq m)$, Chan et al. (1988) also developed index C_{pm}^*, a generalization of C_{pm}, which is defined as:

$$C_{pm}^* = \frac{\min\{D_L, D_U\}}{3\sqrt{\sigma^2 + (\mu - T)^2}},$$

where $D_L = T - LSL, D_U = USL - T$. The index C^*_{pm} reduces to the original index C_{pm} if $T = m$ (processes with symmetric tolerance). Unfortunately, the sampling distribution of the natural estimator of C^*_{pm} is rather complicated.

In attempting to simplify the complication, Greenwich et al. (1995) introduced an index called C_{pp} which is easier to use and analytically tractable. In fact, the index C_{pp} is a simple transformation of the index C^*_{pm}, $C_{pp} = (1/C^*_{pm})^2$, which provides an uncontaminated separation between information concerning the process accuracy and the process precision while such separated information is not available with the index C^*_{pm}. If we denote $D = \min\{D_L, D_U\}/3$, then C_{pp} can be written as:

$$C_{pp} = \frac{(\mu - T)^2}{D^2} + \frac{\sigma^2}{D^2}$$

Some C_{pp} values commonly used as quality requirements in most industry applications are, $C_{pp} = 1.00, 0.56, 0.44, 0.36, \textit{and } 0.25$. A process is called "inadequate" if $C_{pp} > 1.00$, called "marginally capable" if $0.56 < C_{pp} \le 1.00$, called "capable" if $0.44 < C_{pp} \le 0.56$, called "good" if $0.36 < C_{pp} \le 0.44$, called "excellent" if $0.25 < C_{pp} \le 0.36$, and is called "super" if $C_{pp} \le 0.25$.

2. ESTIMATING C_{pp} BASED ON SINGLE SAMPLE

2.1. A Traditional Frequentist Approach

The natural estimator of C_{pp} can be obtained by substituting the sample mean $\bar{X} = \sum_{i=1}^{n} X_i / n$ for μ and the maximum likelihood estimator $S_n^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 / n$ for σ^2 in expression (4), which can be expressed as

$$\hat{C}_{pp} = \frac{(\bar{X} - T)^2}{D^2} + \frac{S_n^2}{D^2} \tag{4}$$

Under the normality assumption, Pearn and Lin (2001) showed that $\hat{C}_{pp}$ is the uniformly minimum variance unbiased estimator (UMVUE) of C_{pp}. Lin (2004) provided maximum values based on the UMVUE $\hat{C}_{pp}$ to develop a reliable decision-making procedure for judging whether or not the process satisfies the preset quality requirement. The probability density function (pdf) can be expressed as (e.g. Pearn and Lin, 2001, 2002):

$$f(x) = \left\{ \sum_{i=0}^{\infty} \frac{(gx)^{i+(n/2)} \exp(-gx)}{x\Gamma\left[i + (n/2)\right]} \times \frac{(\xi/2)^i \exp(-\xi/2)}{\Gamma(i+1)} \right\}, \tag{5}$$

where $g = nD^2/(2\sigma^2), \xi = n(\mu - T)^2/\sigma^2$, and 0 < x < ∞. Recently, Chen et al. (2005) applies the incapability index C_{pp} to develop a graphic evaluation model for measuring supplier quality performance. However, contributions presented above are all based on the traditional frequentist approach.

2.2. A Bayesian Approach

To assess the process capability, Lin (2005) considered the posterior probability Pr{ process is capable | **x** } and proposed a Bayesian approach for assessing process capability by finding a 100p% credible interval, which covers 100p% of the posterior distribution for the incapability index C_{pp}. Compared with the traditional frequentist approach, Bayesian approach has the advantage of providing a statement on the posterior probability that the process is capable under the observed sample data.

Assuming that $\{x_1, x_2, ...x_n\}$ is a random sample taken from $N(\mu, \sigma^2)$, a normal distribution with mean μ and variance σ^2. Adopting the prior $\pi(\mu, \sigma) = 1/\sigma$ and the posterior probability density function $f(\mu, \sigma | x)$ *of* (μ, σ).

$$f(\mu, \sigma | x) = \sqrt{\frac{2n}{\pi}} \frac{\sigma^{-(n+1)}}{\beta^{\alpha}\Gamma(\alpha)} \exp\left\{ -\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2} \right\},$$

where $x = \{x_1, x_2, ..., x_n\}, -\infty < \mu < \infty, 0 < \sigma < \infty, \alpha = (n-1)/2, \beta = 2(nS_n^2)^{-1}$. Given a pre-specified capability level $C_0 > 0$, the posterior probability based on C_{pp} that a process is capable is given as (e.g. Lin, 2005):

$$p = \int_0^{1/t} \frac{\Phi(b_1,(y) + b_2(y)) - \Phi(b_1,(y) + b_2(y))}{\gamma^{\alpha} y^{\alpha+1}\Gamma(\alpha)} \exp\left(-\frac{1}{\gamma y} \right) dy, \tag{6}$$

where Φ is the cumulative distribution function of the standard normal distribution $N(0,1), b_1(y) = \sqrt{2\delta^2/\left[y(1+\delta^2) \right]}, y = (2\sigma^2)/\sum_{i=1}^{n}(X^i - T)^2, \delta = \left| \bar{X} - T \right| / S_n, b_2(y) = \sqrt{n\{[1/(ty)] - 1\}}, t = n\hat{C}_{pp}/(2C_0), \gamma = 1 + \delta^2.$

3. ESTIMATING C_{pp} BASED ON MULTIPLE SAMPLES

3.1. A Traditional Frequentist Approach

In real-world practice, process information is often derived from multiple samples rather then from single sample. For multiple samples of m groups each of size n taken from a stable process, Lin (2006b) considered the following natural estimator of C_{pp} based on multiple samples:

$$\hat{C}_{pp}=\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{(X_{ij}-T)^2}{D^2}=\frac{(\bar{\bar{X}}-T)^2}{D^2}+\frac{S_{mn}^2}{D^2},$$

where $\bar{\bar{X}}=\sum_{i=1}^{m}\bar{X}_i/m,\ \ \bar{X}_i=\sum_{j=1}^{n}X_{ij}/n$ is the i th sample mean, and $S_{mn}^2=\sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij}-\bar{X})^2/(mn)$.

Assuming that the measurements of the characteristic investigated, { X_{i1}, X_{i2}, …, X_{in} }, are chosen randomly from a stable process which follows a normal distribution $N(\mu,\sigma^2)$ for $i=1,2,...,m,$ Lin (2006b) investigated the distributional and inferential properties of $\tilde{C}_{pp}$ Lin (2006b) showed that is the UMVUE of $\tilde{C}_{pp}$ based on multiple samples. Lin (2006b) also derived the r th moment of $\tilde{C}_{pp}$ and constructed upper confidence limits based on the UMVUE $\tilde{C}_{pp}$ to develop a reliable decision-making procedure for judging whether or not the process satisfies the preset quality requirement. The pdf of $\tilde{C}_{pp}$ can be expressed as (e. g. Lin, 2006b):

$$f(x)=\left\{\sum_{i=0}^{\infty}\frac{(g^*x)^{i+(mn/2)}\exp(-g^*x)}{x\Gamma\left[i+(mn/2)\right]}\times\frac{(\xi^*/2)^i\exp(-\xi^*/2)}{\Gamma(i+1)}\right\}, \tag{7}$$

where $g^*=mnD^2/(2\sigma^2),\ \xi^*=mn(\mu-T)^2/\sigma^2,$ and $0<x<\infty$. We note that expression (7) is identical to expression (5) as m = 1. Nevertheless, the sampling distributions of the estimated C_{pp} are rather complicated and intractable as shown in expressions (5) and (7).

3.2. A Bayesian Approach

To assess the process capability based on multiple samples, Lin (2007) considered the posterior probability Pr{ process is capable | **x** } and proposed a Bayesian approach based on multiple samples to evaluate the process capability. Assume that the measurements of the characteristic investigated, $\{x_{i1},x_{i2},...x_{in}\}$, are chosen randomly from a stable process which follows a normal distribution $N(\mu,\sigma^2)$ for i = 1, 2, …, m.

By choosing the prior $\pi(\mu,\sigma)=1/\sigma$, the posterior probability density function $f(\mu,\sigma|x)$ of (μ,σ) based on multiple samples can be expressed as:

$$f(\mu,\sigma|x)=\frac{2\exp\left[(-\sigma^2\beta^*)^{-1}\right]}{\sigma^{mn}\Gamma(a^*)(\beta^*)a^*}\times\sqrt{\frac{mn}{2\pi\sigma^2}}\exp\left[-\left(\frac{mn}{2}\right)\left(\frac{\mu-\bar{\bar{x}}}{\sigma}\right)^2\right],$$

where $x=\{x_{11},x_{12},...,x_{mn}\},-\infty<\mu<\infty,0<\sigma<\infty,0<\sigma<\infty,\ \alpha^*=(mn-1)/2,$

$\beta^* = 2\left[(mn)S_{mn}^2\right]^{-1}$ Given a pre-specified quality requirement $C_0 > 0$, the posterior probability based on C_{pp} with multiple samples can be derived as (e. g. Lin, 2007):

$$p^* = \int_0^{1/t^*} \frac{\Phi(b_1^*(y^*) + b_2^*(y^*)) - \Phi(b_1^*(y^*) - b_2^*(y^*))}{(\gamma^*)^{\alpha^*}(y^*)^{\alpha^*+1}\Gamma(\alpha^*)} \exp(-\frac{1}{\gamma^* y^*}) dy^*, \tag{8}$$

where Φ is the cumulative distribution function of the standard normal distribution

$$N(0,1), b_1^*(y^*) = \sqrt{2\delta^{2*}/\left[y^*(1(\delta^*)^2)\right]}, y^* = 2\sigma^2 / \left[\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij}-T)^2\right],$$

$$\delta^* = \left|\overline{\overline{X}}-T\right| / S_{mn}, \quad b_2^*(y^*)\sqrt{mn\left\{\left[1/(t^*y^*)\right]-1\right\}}, \gamma^* = 1+\delta^{*2}, t^* = mn\tilde{C}_{pp}/(2C_o).$$

Note that expression (8) can be reduced to expression (6) as m = 1.

In our Bayesian approach based on multiple samples, we say that a process with symmetric production tolerance is capable if all the points fall within this credible interval are less than a pre-specified value of C_0. When this occurs, we have Pr{ process is capable $|x$ } > p^*. Therefore, to test whether or not a process is capable (with capability level C_0 and credible level p^*), we only need to check whether or not $\tilde{C}_{pp} < C_0 C^*(p^*)$.

For the well-centered case in which $\mu = T$, the formula for $C_{pp} = \left[(\mu - T)/D\right]^2 + (\sigma + D)^2$ reduce to $C_{ip} = (\sigma / D)^2$ and we could use the UMVUE $C_{ip} = (S/D)^2$ proposed by Lin (2006b), where $S^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{\overline{X}})^2/(mn-1)$ is unbiased for σ^2. We note that $\left[(mn-1)/C_{ip}\right]\tilde{C}_{ip} = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{\overline{X}})^2/\sigma^2$ is distributed as $\chi^2(mn-1)$, a chi-squared distribution with $(mn-1)$ degrees of freedom. The posterior probability for a well-centered process is capable is given as $p^* = \left\{C_{ip} < C_0 \left|x\right.\right\} = \Pr\left\{\chi^2(mn-1) > (mn-1)\tilde{C}_{ip}/C_0\right\}$. Thus, to compute p^*, we need only check the commonly available chi-squared tables for the posterior probability p^*. If p^* is greater than a desirable level, say 95%, then we may claim that the process is capable (in a Bayesian sense) with 95% confidence.

4. A CONTAMINATED MODEL

4.1.The Joint Distribution of k Contaminated Normal Processes

The contamination model provides a rich class of distributions that can be used in modeling populations with combined (mixed) characteristics. The contamination model is useful, particularly, for cases with multiple manufacturing processes where the equipment, or workmanship may not be identical in precision and consistency for each manufacturing process, or cases where multiple suppliers are involved in providing raw materials for the manufacturing. Such situations often result in productions with inconsistent precision in quality characteristic, and using the contaminated model to characterize the process would be appropriate. We consider a contamination model of k normal populations, having probability density function:

$$fx(x) = \sum_{j=1}^{k} p_j \phi(x; \mu_j, \sigma), \tag{9}$$

where $0 \le p_j \le 1$, and $\phi(x;\mu_j,\sigma) = (1/\sqrt{2\pi\sigma})\exp\left[-(x-\mu_j)^2/2\sigma^2\right]$. We note that random samples of size n from a population with probability density function defined as $f(x)$ can be regarded as mixtures of random samples with $N_1, N_2, ..., N_k$ individual observations from populations with probability density functions $\phi(x;\mu_1,\sigma), ; \phi(x;\mu_2,\sigma), \ldots,$ and $\phi(x;\mu_k,\sigma)$, where $N_1, N_2, ..., N_k$ have the following joint distribution with $0 \le p_i \le 1$, $\sum_{j=1}^{k} p_j = 1$ and $\sum_{j=1}^{k} n_j = n$.

$$P(N=n) = P\left\{\bigcap_{j=1}^{k}(N_j = n_j)\right\} = \frac{n!}{n_1!n_2!...n_k!} p_1^{n_1} p_2^{n_2} ... p_k^{n_k} \tag{10}$$

4.2. Estimating C_{pp} Based on k Contaminated Normal Processes

Suppose that $X_1, X_2, ...X_n$ represent the sample values with n_j observations of X`s from, $\phi(x;\mu_j,\sigma)$, $j = 1,2,...,k.$ Then, given $N = n$ the conditional distribution of C_{pp} is that of $\left[(\sigma/D)^2/n\right]\times$ non-central chi- squared with n degrees of freedom and non-centrality parameter

$$\tau(n) = \sum_{j=1}^{k} \frac{n_j(\mu_j - T)^2}{\sigma^2} \tag{11}$$

Given $N = n$ the conditional r-th moment of $\hat{C}_{pp}$ can be calculated as

$$E\left(\hat{C}_{pp}^r \,|N = n\right) = \left(\frac{\sigma^2}{nD^2}\right)^r E\left[\chi_n^2(\tau(n))\right]^r$$

$$= \sum_{j=0}^{\infty} \frac{\Gamma\left(j + \frac{n}{2} + r\right)}{\Gamma\left(j + \frac{n}{2}\right)} \left(\frac{2\sigma^2}{nD^2}\right)^p \frac{\left[\tau(n)/2\right]^j exp\left[-\tau(n)/2\right]}{\Gamma(j+1)}$$

Hence, the r-th moment of is $\hat{C}_{pp}$

$$E(\hat{C}_{pp}^r) = \sum_{\underset{\sim}{n}} E(\tilde{C}_{pp}^r \,|\underset{\sim}{N} = \underset{\sim}{n}) P(\underset{\sim}{N} = \underset{\sim}{n}) = \Psi(\frac{\sigma^2}{D^2})^r, \tag{12}$$

$$\Psi = \sum_{\underset{\sim}{n}} \sum_{j=0}^{\infty} \frac{\Gamma(j+\frac{n}{2}+r)}{\Gamma(j+\frac{n}{2})} \left(\frac{2}{n}\right)^r P(\underset{\sim}{N}=\underset{\sim}{n}) \frac{\left[\tau(\underset{\sim}{n})/2\right] j \exp\left[-\tau(\underset{\sim}{n}/2)\right]}{\Gamma(j+1)}$$

If $p_1 = 1$ (no contamination in this case), then $\tau(\underset{\sim}{n})$ reduces to $n(\mu - T)^2/\sigma^2$ and Ψ reduces to

$$\Psi = \sum_{j=0}^{\infty} \frac{\Gamma(j+\frac{n}{2}+r)}{\Gamma(j+\frac{n}{2})} \left(\frac{2}{n}\right)^r \frac{\left[\tau(n)/2\right]^j \exp\left[-\tau(n/2)\right]}{\Gamma(j+1)}$$

Therefore, the r-th moment of $\hat{C}_{pp}$ can be simplified to

$$E(\hat{C}_{pp}^r) = \Psi \left(\frac{\sigma^2}{D^2}\right)^r, = \sum_{j=0}^{\infty} \frac{\Gamma(j+\frac{n}{2}+r)}{\Gamma(j+\frac{n}{2})} \left(\frac{2\sigma^2}{nD^2}\right)^r \frac{\left[\tau(n)/2\right]^j \exp\left[-\tau(n/2)\right]}{\Gamma(j+1)}$$

The result is identical to that obtained by Pearn and Lin (2001) for the normal case.

5. CONCLUSIONS

Existing developments and applications of the incapability index have focused on single normal process. In this paper, investigations based on contaminated normal processes of the estimated incapability index were considered. The exact sampling distributions and r-th moments of the estimated index were derived. The proposed contaminated model can provide an efficient alternative to the traditional single normal process approach in assessing process performance.

Acknowledgment

The author would like to thank the anonymous referees for their helpful comments, which significantly improved the paper. This research was partially supported by National Science Council of the Republic of China (NSC-94-2213-E-151-026).

REFERENCES

[1] Boyles, R.A., "The Taguchi capability index", *Journal of Quality Technology*, 23 (1991) 17-26.

[2] Chan, L.K., Cheng, S.W., and Spiring, F.A., "A new measure of process capability: C_{pm}", *Journal of Quality Technology*, 20 (1988) 162-175.

[3] Chen, K.L., Chen, K.S., and Li, R.K., "Suppliers capability and price analysis chart", *International Journal of Production Economics*, 98 (2005) 315- 327.

[4] Chou, Y.M., and Owen, D.B., "On the distributions of the estimated process capability indices", *Communication in Statistics - Theory and Methods*, 18 (1989) 4549-4560.

[5] Greenwich, M., and Jahr-Schaffrath, B.L., "A process incapability index", *International Journal of Quality & Reliability Management*, 12 (1995) 58-71.

[6] Johnson, T., "The relationship of C_{pm} to squared error loss", *Journal of Quality Technology*, 24 (1992) 215-221.
[7] Kane, V.E., "Process capability indices", *Journal of Quality Technology*, 18 (1986) 41-52.
[8] Kotz, S., Pearn, W.L., and Johnson, N.L., "Some process capability indices are more reliable than one might think", *Applied Statistics*, 42 (1993) 55- 62.
[9] Lin, G.H., "Upper limits of the estimated incapability index: A practical application on the reliability assessment of the printed circuit boards", *The International Journal of Advanced Manufacturing Technology*, 24 (2004) 841-846.
[10] Lin, G.H., "Process reliability assessment with a Bayesian approach", *The International Journal of Advanced Manufacturing Technology*, 25 (2005) 392-395.
[11] Lin, G.H., "A random interval estimation on the estimated process accuracy index", *The International Journal of Advanced Manufacturing Technology*, 27 (2006a) 969-974.
[12] Lin, G.H., "Assessing process performance with incapability index based on multiple samples", *The International Journal of Advanced Manufacturing Technology*, 27 (2006b) 715-719.
[13] Lin, G.H., "A Bayesian approach based on multiple samples for measuring process performance with incapability index", *International Journal of Production Economics*, 106 (2007) 506-512.
[14] Pearn, W.L., Kotz, S., and Johnson, N.L., "Distributional and inferential properties of process capability indices", *Journal of Quality Technology*, 24 (1992) 216-231.
[15] Pearn, W.L., and Lin, G.H., "On the reliability of the estimated process incapability index", *Quality & Reliability Engineering International*, 17 (2001) 279–290.
[16] Pearn, W.L., and Lin, G.H., "Estimated incapability index: Reliability and decision making with sample information", *Quality & Reliability Engineering International*, 18 (2002) 141-147.
[17] Pearn, W.L., Lin, G.H., and Chen, K.S., "Distributional and inferential properties of the process accuracy and process precision indices", *Communications in Statistics – Theory and Methods*, 27 (1998) 985– 1000.