Pak. J. Statist.
2017 Vol. 33(6), 411-418

A MODIFIED APPROACH FOR ESTIMATING PROCESS CAPABILITY INDICES USING IMPROVED ESTIMATORS

Senem Şahan Vahaplar[1§] and **Özlem Ege Oruç**[2]
Department of Statistics, Dokuz Eylül University, Türkiye
Email: [1]senem.sahan@deu.edu.tr; [2]ozlem.ege@deu.edu.tr
[§]Corresponding author

ABSTRACT

In this paper, new estimators for process capability indices have been proposed using improved estimators of population mean and variance. The proposed and classical indices are compared with respect to mean square error. A numerical example is considered to illustrate the proposed indices and to judge the merits of the proposed estimators.

KEYWORDS

Statistical process control, Process capability indices, Improved estimation, Mean square error.

1. INTRODUCTION

Process capability analysis has been studied by many researchers in both theoretical and applied fields. Process capability indices quantify how much the process is able to produce required items. Kotz and Lovelace (1998) and Wu et al. (2009) define the general idea of process capability indices as comparing "what the process should do" with "what the process is actually doing". Kane (1986) states that these indices relate to the natural tolerance and engineering specifications. Details of these indices are given by Kotz and Johnson (1993) and Kotz and Lovelace (1998).

In statistical process control, two assumptions have to be satisfied for conducting a process capability analysis; *(i)* the process should be in a state of statistical control, *(ii)* the quality characteristic should be normally distributed (Montgomery, 2009; Stoumbos, 2002; Ali et al., 2015; Vännman and Kotz, 1995; Kane, 1986; Nagata, 1995; Maiti et al., 2010). Process capability indices give reliable results when these assumptions are satisfied.

Improved estimation plays an important role in statistical inference. The main purpose of this estimation technique is to examine the conditions under which biased estimators can lead to an improvement over the conventional unbiased procedures. Given additional information such as coefficient of variation, kurtosis or skewness, the problem has been studied extensively by Searls (1964), Singh et al. (1973), Bibby and Toutenburg (1974) and Arnholt and Hebert (1995). Compared to the usual procedures, improved estimation provides more efficient estimators with respect to mean square error.

The aim of this paper is to introduce new estimators of process capability indices by using improved estimators of population mean and variance. It should be noted that improved estimators of population parameters are generally biased estimators but they have lower mean square error (MSE).So, process variability is considered to be reduced since proposed estimators are biased with lower mean square error with regard to the classical estimators of process capability indices.

The paper is organized as follows. First, we provide an overview of process capability indices and improved estimators of population mean and variance. Then we propose improved estimators of process capability indices. Finally, we compare them with the classical estimators of process capability indices and evaluate the results.

2. PROCESS CAPABILITY INDICES

C_p seems to be the first process capability index proposed in the literature (Juran, 1974; Kane, 1986). It assesses whether the natural tolerance of a process is within the specification limits. This means comparing allowable process spread with the actual process spread as given in equation (1),

$$C_p = \frac{allowable\ process\ spread}{actual\ process\ spread} = \frac{USL - LSL}{6\sigma} \tag{1}$$

where USL and LSL are upper and lower specification limits, respectively (Juran, 1974; Wu et al., 2009; Kane, 1986).

C_p is stated to be inadequate for off-center processes. For an off-center process; the process mean is not equal to the target value (T). For such processes, C_{pk} index provides a good alternative. It takes both the process spread and the departure of the process mean from the target value into consideration.

$$C_{pk} = \min\left\{\frac{USL - \mu}{3\sigma} \quad , \quad \frac{\mu - LSL}{3\sigma}\right\} = \frac{d - |\mu - m|}{3\sigma} \tag{2}$$

d is the half specification width defined in (3) and m is the midpoint between upper and lower specification limits as in (4).

$$d = \frac{USL - LSL}{2} \tag{3}$$

$$m = \frac{USL + LSL}{2} \tag{4}$$

Despite being more explanatory than C_p index, C_{pk} still has a disadvantage. It cannot provide information about the location of process mean in the tolerance interval (LSL, USL). Also, C_p and C_{pk} are not related to the cost of failing to meet the target requirement of the customers. For this reason, Hsiang and Taguchi (1985) and Chan et al. (1988) proposed C_{pm} index,

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - T)^2}} = \frac{d}{3\tau} \tag{5}$$

where τ is a measure of the average product deviation from the target value.

$$\tau = E\left[(X - T)^2\right] = \sqrt{\sigma^2 + (\mu - T)^2} \tag{6}$$

Pearn et al. (1992) introduced C_{pmk} index having all the properties of previous indices. Two important components of this index can be summarized as *(i)* variation relative to the process mean (σ^2) and *(ii)* deviation of the process mean from the target $(\mu - T)^2$. Therefore, C_{pmk} is more sensitive than previous indices (Vannman, 1995).

$$C_{pmk} = \min\left\{\frac{USL - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}} \quad , \quad \frac{\mu - LSL}{3\sqrt{\sigma^2 + (\mu - T)^2}}\right\} = \frac{d - |\mu - m|}{3\sqrt{\sigma^2 + (\mu - T)^2}} \tag{7}$$

3. PROPOSED ESTIMATORS OF PROCESS CAPABILITY INDICES

In this section, alternative estimators of four process capability indices are proposed. It is assumed that the process is under statistical control and approximately normally distributed.

Let $x_1, x_2, \ldots, x_n$ be a random sample from a normally distributed process with unknown mean μ and variance σ^2. It is well known that the unbiased estimators of μ and σ^2 are $\bar{x}$ and s^2, respectively.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{8}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{9}$$

Searls (1964) defined an improved estimator of the population mean as given,

$$\bar{x}^* = \frac{1}{n + v^2}\sum_{i=1}^{n} x_i \tag{10}$$

where v is the known coefficient of variation of the distribution. It can be seen that $MSE\left(\bar{x}^*\right)$ is always less than $MSE\left(\bar{x}\right)$.

Singh et al. (1973) defined an improved estimator of the population variance as,

$$s^{2*} = \frac{n}{n^2 - 2n + 3 + \beta_2 (n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{11}$$

where β_2, the coefficient of kurtosis, is known as a priori information. It is well known that the coefficient of kurtosis can be calculated through equation (12) by using the fourth moment around the origin. It can be seen that $MSE\left(s^{2*}\right)$ is always less than $MSE\left(s^2\right)$.

$$\beta_2 = \frac{\mu_4}{\sigma^4} \tag{12}$$

Using the improved estimators of population mean and variance, the estimators of process capability indices are obtained as follows.

$$\hat{C}_p^{\,*} = \frac{USL - LSL}{6\hat{\sigma}} = \frac{USL - LSL}{6\sqrt{\dfrac{n}{n^2 - 2n + 3 + \beta_2 (n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2}} \tag{13}$$

Here, it was formerly stated that the process is normally distributed. Besides, β_2 is assumed to be known as a prior information about the process. As the value of coefficient of kurtosis is 3 for normal distribution, it is substituted in equation (13) and (14) is obtained.

$$\hat{C}_p^{\,*} = \frac{USL - LSL}{6\sqrt{\dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n+1}}} \tag{14}$$

The estimator of C_{pk} is obtained as

$$\hat{C}_{pk}^{\,*} = \frac{d - |\hat{\mu} - m|}{3\hat{\sigma}} = \frac{d - \left|\dfrac{\sum x_i}{n + v^2} - m\right|}{3\sqrt{\dfrac{\sum (x_i - \bar{x})^2}{n+1}}} \tag{15}$$

where v is the coefficient of variation for normal distribution which is $v = \sigma/\mu$.

The estimator of C_{pm} index is given below.

$$\hat{C}_{pm}^{\,*} = \frac{USL - LSL}{6\sqrt{\hat{\sigma}^2 + (\hat{\mu} - T)^2}} = \frac{USL - LSL}{6\sqrt{\dfrac{\sum (x_i - \bar{x})^2}{n+1} + \left(\dfrac{\sum x_i}{n + v^2} - T\right)^2}} \tag{16}$$

Finally, the estimator of C_{pmk} is given in equation (17).

$$\hat{C}_{pmk}^{\,*} = \frac{d - |\hat{\mu} - m|}{6\sqrt{\hat{\sigma}^2 + (\hat{\mu} - T)^2}} = \frac{d - \left| \frac{\sum x_i}{n + v^2} - m \right|}{6\sqrt{\frac{\sum (x_i - \bar{x})^2}{n+1} + \left(\frac{\sum x_i}{n + v^2} - T \right)^2}} \tag{17}$$

4. COMPARISON OF CLASSICAL AND PROPOSED INDICES

In order to research the behavior of classical and proposed indices, an application is realized using Minitab. For this application, a sample data set from Montgomery (2009) is used. Data set consists of flow width measurement data of the resistance in semiconductor manufacturing. The target value of flow width is 1,5 microns. Also, lower and upper specification limits are 1 and 2 microns, respectively. Histogram of data with USL, LSL and target value is given in Figure 1.

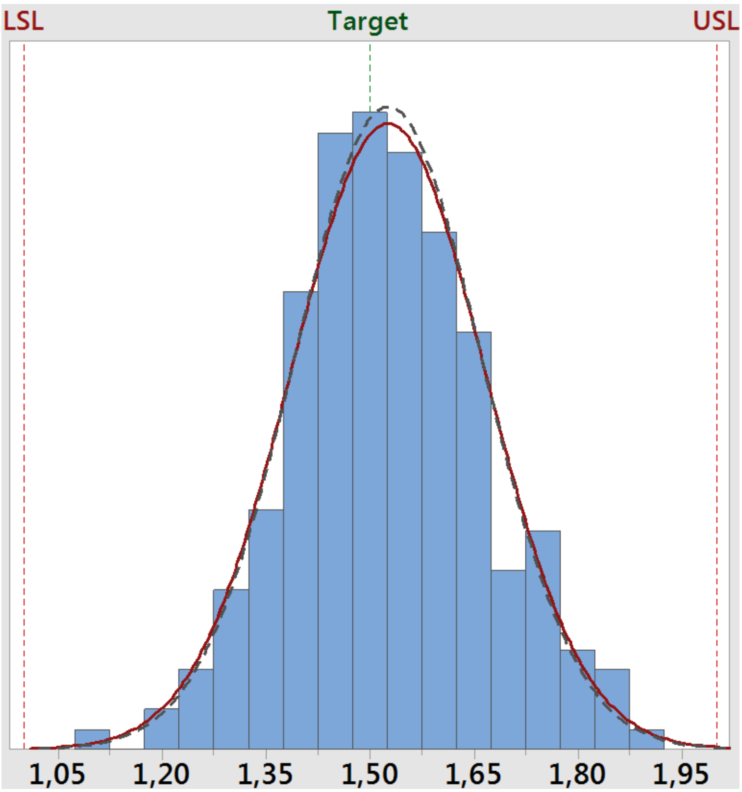

Figure 1: Histogram of Flow Width Data with LSL, USL and Target Value

As mentioned before, there are two assumptions of process capability analysis. So we need to check whether the assumptions are satisfied or not. Anderson-Darling normality test shows that the data are distributed normally. The output of the test is given in Figure 2.

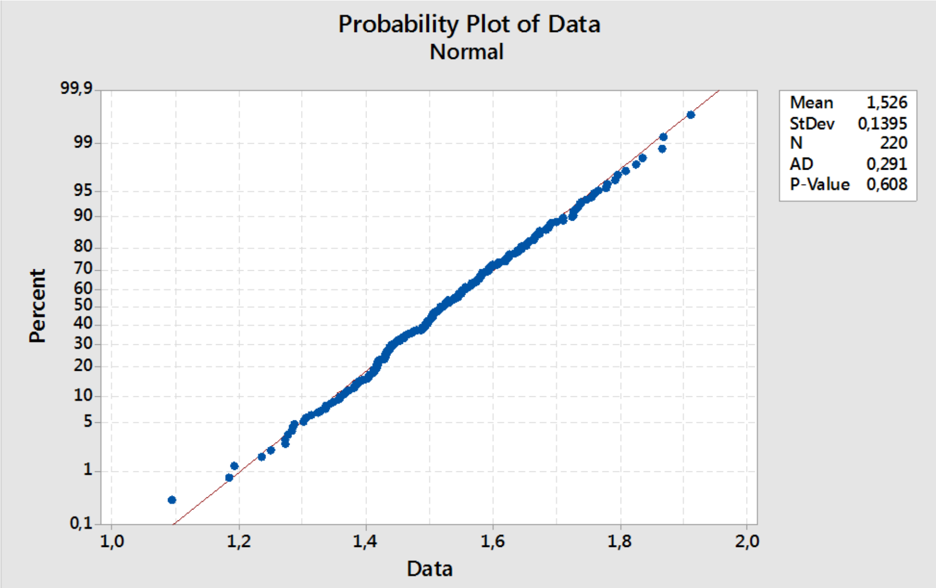

Figure 2: Output of the Normality Test for Flow Width Data

$\bar{x}$ & s charts in Figure 3 show that none of the observations fall beyond the control limits and the process is in statistical control.

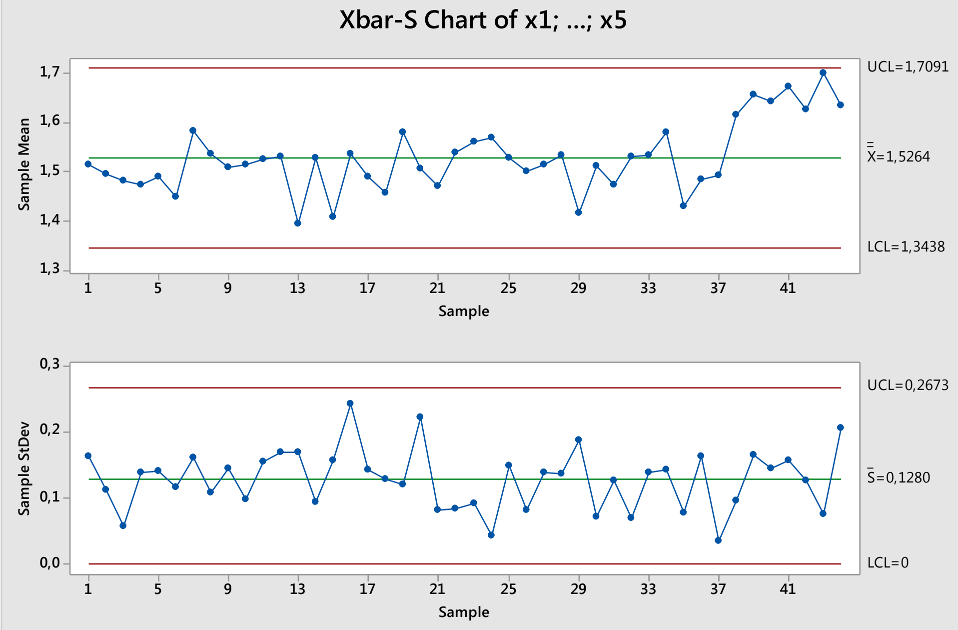

Figure 3: $\overline{x}$ **& s Control Charts for Flow Width Data**

Both assumptions are satisfied so the estimates of indices are calculated as in Table 1. It can be seen that proposed estimates for all four indices are greater than their classical estimates. MSE of proposed estimators are lower than classical estimators. This means that the variability of proposed estimators are lower. So the estimations obtained by proposed estimators can be more reliable than the classical estimators because of lower variability.

Table 1
Classical and Proposed Estimates of Process Capability Indices for Flow Width Data

Process Capability Index	Classical Estimate	Proposed Estimate
C_p	1.1947	1.2002
C_{pk}	1.1316	1.1369
C_{pm}	1.1738	1.1791
C_{pmk}	1.1118	1.1169

5. CONCLUDING REMARKS

In this paper, our focus has been introducing new estimators of process capability indices by using improved estimation of population mean and variance. The proposed estimators are biased but they have lower mean squared error.

To compare the estimators, a sample data set is used. First, normality test is realized to see if the process is normally distributed or not. Then, quality control charts are used to see whether the process is in a state of statistical control. Finally, classical and proposed estimates of C_p, C_{pk}, C_{pm} and C_{pmk} indices are obtained. According to the results, proposed estimates for all four indices are greater than their classical estimates.

REFERENCES

1. Ali S., Aslam, M., Abbas, N., Kazmi, S.M.A. and Hasan, T. (2015). On process capability and system availability analysis of the inverse Rayleigh distribution. *Pakistan Journal of Statistics and Operation Research*, XI(1), 53-66.
2. Chan, L.K., Cheng, S.W. and Spiring, F.A. (1988). A new measure of process capability C_{pm}. *Journal of Quality Technology*, 20(3), 162-175.
3. Hsiang, T.C. and Taguchi, G. (1985). A tutorial on quality control and assurance - The Taguchi methods. In: *ASA Annual Meeting*, Las Vegas.
4. Juran, J.M. (1974). *Quality Control Handbook, 3rd edition*. McGraw-Hill, New York.
5. Kane, V.E. (1986). Process capability indices. *Journal of Quality Technology*, 18(1), 41-52.
6. Kotz, S. and Johnson, N.L. (1993). *Process Capability Indices*. Chapman & Hall, London.
7. Kotz, S. and Lovelace, C. (1998). *Process Capability Indices in Theory and Practice*. Arnold, London.
8. Maiti, S.S., Saha, M. and Nanda, A.K. (2010). On generalizing process capability indices. *Quality Technology & Quantitative Management*, 7(3), 279-300.
9. Montgomery, D.C. (2009). *Introduction to Statistical Quality Control, 6th Edition*. Wiley, United States of America.
10. Nagata, Y. (1995). 永田靖 - 岡山大学経済学会雑誌, 26(3-4), 325-339.
11. Pearn, W.L., Kotz, S. and Johnson, N.L. (1992). Distributional and inferential properties of process capability indices. *Journal of Quality Technology*, 24(4), 216-231.
12. Searls, D.T. (1964). The utilizing of a known coefficient of variation in estimation procedure. *Journal of American Statistical Association*, 59, 1225-1226.
13. Singh, J., Pandey, B.N. and Hirano, K. (1973). On the utilizing of a known coefficient of variation of kurtosis in estimation procedure of variance. *Annals of the Institute of Statistical Mathematics*, 25, 51-55.
14. Stoumbos, Z.G. (2002). Process capability indices: overview and extensions. *Nonlinear Analysis: Real World Applications*, 3, 191-210.
15. Vännman, K. and Kotz, S. (1995). A superstructure of capability indices - Distributional properties and implications. *Scandinavian Journal of Statistics*, 22(4), 477-491.
16. Vännman, K. (1995). A unified approach to capability indices. *Statistica Sinica*, 5, 805-820.
17. Wu, A.W., Pearn, W.L. and Kotz, S. (2009). An overview of theory and practice on process capability indices for quality assurance. *International Journal of Production Economics*, 117, 338-359.