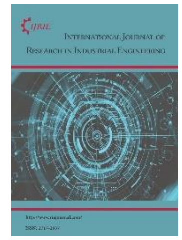
Int. J. Res. Ind. Eng. Vol. 9, No. 3 (2020) 286–303
DOI:10.22105/riej.2020.237520.1137

International Journal of Research in Industrial Engineering

www.riejournal.com



A Note on the Generalized Indices of Process Capability

S. Kumar Gauri[1,*], S. Pal[2]

[1] Statistical Quality Control and Operations Research Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata-700108, India.
[2] Statistical Quality Control and Operations Research Unit, Indian Statistical Institute, 110, Nelson Manickam Road, Chennai- 600029, India.

([*]Corresponding author Email Address: susantagauri@hotmail.com)

A B S T R A C T

Process capability indices are widely used to assess whether the outputs of an in-control process meet the specifications. The commonly used indices are C_p, C_{pu}, C_{pl} and C_{pk}. In most applications, the quality characteristics are assumed to follow normal distribution. But, in practice, many quality characteristics, e.g. count data, proportion defective etc. follow Poisson or binomial distributions, and these characteristics usually have one-sided specification limit. In these cases, computations of C_{pu} or C_{pl} using the standard formula is inappropriate. In order to alleviate the problem, some generalized indices (e.g. C index, C_f index, C_{pc} index and C_{py} index) are proposed in literature. The variant of these indices for one-sided specification are C_u and C_l, C_{fu} and C_{fl}, C_{pcu} and C_{pcl}, C_{pyu} and C_{pyl} respectively. All these indices can be computed in any process regardless of whether the quality characteristics are discrete or continuous. However, the same value for different generalized indices and C_{pu} or C_{pl} signifies different capabilities for a process and this poses difficulties in interpreting the estimates of the generalized indices. In this study, the relative goodness of the generalized indices is quantifying capability of a process is assessed. It is found that only C_u or C_l gives proper assessment about the capability of a process. All other generalized indices give a false impression about the capability of a process and thus usages of those indices should be avoided. The results of analysis of multiple case study data taken from Poisson and binomial processes validate the above findings.

Keywords: Generalized PCI, Goodness of generalized PCI, normal process, binomial process, Poisson process.

Article history: *Received: 17 March 2020* *Reviewed: 26 April 2020* *Revised: 12 June 2020* *Accepted: 21 July 2020*

Kumar Gauri, S., & Pal, S. (2020). A note on the generalized indices of process capability. *International journal of research in industrial engineering*, *9*(3), 286-303.

1. Introduction

Process capability refers to the ability of a process in producing outputs according to specified requirements. Therefore, process capability analysis is of great importance to the design engineers, process managers, vendors as well as customers. Process Capability Index (PCI) is widely used for quantifying the capability of a process. The basic process capability indices are $C_p = (USL - LSL)/6\sigma$ and $C_{pk} = Min[(USL - \mu)/3\sigma, (\mu - LSL)/3\sigma]$, where USL and LSL are the upper and lower specification limits; μ is the population mean, and σ is the population standard deviation. When there is only USL or only LSL for a quality characteristic, then the process capability indices are defined as $C_{pu} = (USL - \mu)/3\sigma$ and $C_{pl} = (\mu - LSL)/3\sigma$ respectively. The most commonly used other indices are C_{pm}, and C_{pmk} [1]. Computation of all these indices requires assumption that the quality characteristic follows normal distribution implying that the characteristic is a continuous variable. The details about these indices are available in Kane [2], Kotz and Johnson [3], English and Taylor [4], Kotz and Johnson [5], Chen et al. [6], Wu et al. [7], Yum and Kim [8] and Chen et al. [9]. The applications (generalization) of these indices for continuous but non-normal variables are suggested by Clements [10], Pearn and Chen [11], Shore [12], Chen [13] and Goswami and Dutta [14].

In practice, many quality characteristics in manufacturing and service set ups are attribute in nature. For example, number of defects in 100 square metre cloths, number of customers served per hour, proportion of improperly sealed orange juice can, proportion of defective purchase orders, proportion of successful transactions etc. The attribute data are typically obtained by counting the number of occurrences of some condition (e.g. defect, error etc.) in an inspection unit or by counting number of defective units (r) within a given number of sample units(n), and so, these data are discrete in nature. It is well established that attribute data usually follow Poisson or binomial distribution. Therefore, standard formulas (that are developed for normal processes) cannot be used for computation of capability indices of a process involving such characteristics.

In order to alleviate the problem, in recent past, some alternative indices are proposed in literature. Borges and Ho [15] proposed C index , Yeh and Bhattacharya [16] proposed C_f index, Perakis and Xekalaki [17-18] presented C_{pc} index and Maiti et al. [19] suggested the C_{py} index for measuring capability of a process in producing outputs within specification limits. In case of unilateral specification, the equivalent generalized indices for C_{pu} and C_{pl} can be represented as C_u and C_l, C_{fu} and C_{fl}, C_{pcu} and C_{pcl}, and C_{pyu} and C_{pyl} respectively. It may be worth to mention here that Poisson or binomial data usually have one-sided specification limit only. For example, number of defects in a printed circuit board (a Poisson variable) or proportion of fraction defective in a lot of products (a binomial variable) is desired to be less than a specified USL. The capability of the concerned processes can be judged by measuring one-sided generalized indices like C_u, C_{fu}, C_{pcu} and C_{pyu}. Similarly, number of customers served per hour (a Poisson variable) or proportion of fraction conforming in a lot of products (a binomial variable) is desired to be more than a specified LSL, and capability of the concerned processes can be assessed by measuring one-sided generalized indices like C_l, C_{fl}, C_{pcl} and C_{pyl}.

The indices C_{fu} (or C_{fl}), C_{pcu} (or C_{pcl}) and C_{pyu} (or C_{pyl}) are computed as the ratio of two probabilities, and thus these indices can be computed for both continuous as well as discrete quality characteristics, and no assumption is required on the distributions of these quality characteristics. On the other hand, C_u (or C_l) is computed by mapping the expected proportion of nonconformance of a characteristic above USL (or below LSL) to the Z-score in the right side of standard normal distribution. That is, C_u (or C_l) responds to changes in the nonconforming region and not to changes in the distribution of the observed quality characteristic. Consequently, computations of all these indices for one-sided specification are feasible in any process regardless of whether the quality characteristics are discrete or continuous and their probability distributions. This is why these indices are called as generalized indices.

It is important to note that since introduction of the concept of process capability index, conventionally the indices C_p, C_{pu} and C_{pl} are estimated from normal processes to facilitate better decision making in product and process management. By virtue of the relationships 1) $P(LSL \leq X \leq USL) = 2\Phi(3 \times \hat{C}_p) - 1$, 2) $P(X \leq USL) = P\left(z \leq 3 \times \frac{USL-\hat{\mu}}{3\hat{\sigma}}\right) = \Phi(3 \times \hat{C}_{pu})$ and 3) $P(X \geq LSL) = 1 - P(X < LSL) = 1 - P\left(z \leq 3 \times \frac{LSL-\hat{\mu}}{3\hat{\sigma}}\right) = 1 - \Phi(-3 \times \hat{C}_{pl})$, the analysts/users of the indices can easily assess/predict the expected proportion of conforming products in the process outputs based on an estimated index value. For example, $\hat{C}_{pu} = 0.5$ implies that the process is capable of producing 93.319% conforming products with respect to USL, $\hat{C}_{pu} = 1$ implies that the process is capable of producing 99.865% conforming products and $\hat{C}_{pu} = 1.3$ implies that the process is capable of producing 99.995% conforming products. Over the years, process managers, engineers and other decision makers have become accustomed to relate the estimates of process capability indices and the expected proportion of product conformance to specifications in this way. Accordingly, general thumb rule is being followed among the users of the indices that the capability of a process is good if $\hat{C}_{pu} \geq 1$ and the capability is very good if $\hat{C}_{pu} \geq 1.33$.

The main advantage of generalized indices is that these indices can be computed in any process regardless of whether the quality characteristics are discrete or continuous. However, the estimates of different generalized indices obtained from the same process usually differ widely. Also equal value of a generalized index and a conventional index signifies different capabilities of the processes in meeting the specifications. All these pose difficulties in interpreting the estimates of the generalized indices. Ideally, the interpretations of the values of generalized indices and interpretations of values of conventional process capability indices should match as closely as possible. Otherwise, the users (process managers, design engineers, vendors, customers etc.) of the generalized indices may unknowingly get a false impression about the capability of the concerned process, which may lead him/her to erroneous decision making. Consequently, product and process management may become inefficient. These shortcomings of the generalized indices motivate us to carry out the current study of assessing relative goodness of different generalized indices in quantifying the capability of a process. The basic premise of the study is that a generalized index may be considered better in quantifying the capability of a process if the interpretations of its values match to the interpretations of the values of the relevant conventional process capability index more closely than others.

In this article, the relative goodness of different generalized indices with respect to USL is assessed. The article is organized as follows: The methods for computation of different generalized indices are described in Section 2. The procedure for assessment of the relative goodness of different generalized indices is described in Section 3. Analysis of multiple case study data on Poisson and binomial processes, and related results are presented in Section 4. Some issues of the generalized indices are discussed in Section 5. Section 6 concludes the paper.

2. Generalized Indices for Process Capability

The generalized indices for process capability proposed by Borges and Ho [15], Yeh and Bhattacharya [16], Perakis and Xekalaki [17-18] and Maiti et al. [19] and their variants for unilateral specification are presented in the following subsections.

2.1. C Index

Borges and Ho [15] suggested a new measure of process capability, called C index, which has one-to-one correspondence (mapping) between the proportion of nonconformance and Z-value of the standard normal distribution. In this method, the expected proportion of nonconformance of a characteristic above USL (or below LSL) is mapped to the Z-score in the right side of standard normal distribution, and 1/3rd of this Z-score is considered as the measure of the process capability with respect to USL (or LSL) and it is denoted as C_u (or C_l).

For computation of estimates of C_u and C_l for count data from a Poisson process, we need to compute first expected proportion of nonconforming units with respect to USL ($\widehat{PNU}_U$) and expected proportion of nonconforming units with respect to LSL ($\widehat{PNU}_L$) respectively. For convenience, let us consider that a single unit of product represents an inspection unit. Suppose number of occurrences of the events (e.g. defects or errors) is observed in each of the m units collected from a stable process. Let the random variable C denotes the number of occurrences of the event in a unit, and c_i is the number of events occurred in the i^{th} unit ($i = 1,2,3 \dots, m$). Then, the unknown parameter λ can be estimated as $\hat{\lambda} = \bar{c} = \sum_{i=1}^{m} c_i/m$, and the values of $\widehat{PNU}_U$ and $\widehat{PNU}_L$can be obtained as follows:

$$\widehat{PNU}_U = P\{c > c_U\} = 1 - P\{c \le c_U\} = 1 - \sum_{c=0}^{c_U} e^{-\bar{c}}(\bar{c})^c/c!\,, \tag{1}$$

$$\widehat{PNU}_L = P\{c < c_L\} = \sum_{c=0}^{c_L-1} e^{-\bar{c}}(\bar{c})^c/c!\,. \tag{2}$$

where c_U is the USL for occurrences of a Smaller-The-Better (STB) type event and c_L is the LSL for occurrences of a Larger-The Better (LTB) type event in a unit. The quantity $\widehat{PNU}_U$ gives the expected proportion of units where each unit will contain more than the specified c_U number of events. The estimates of C_u and C_l for count data from a Poisson process are then obtained as follows:

$$\hat{C}_u = (1/3) \times \Phi^{-1}\left(1 - \widehat{PNU}_U\right). \quad (3)$$

$$\hat{C}_l = (1/3) \times \Phi^{-1}\left(1 - \widehat{PNU}_L\right). \quad (4)$$

On the other hand, for computation of estimates of C_u and C_l from a binomial process, we need to compute first expected proportion of nonconforming lots with respect to USL ($\widehat{PNL}_U$) and expected proportion of nonconforming lots with respect to LSL ($\widehat{PNL}_L$) respectively. Let a production process is operating in a stable manner, such that the probability that any unit will be nonconforming (or conforming) to specification is p and successive units produced are independent. Suppose a random sample of n units of product is selected from the process. If the random variable D denotes the number of units of product that are nonconforming to the standard, then D has a binomial distribution with parameters n and p. The cumulative distribution function of sample fraction nonconforming, $f = d/n$ can be obtained by using the binomial distribution as

$$P\{f \leq a\} = P\left\{\frac{d}{n} \leq a\right\} = P\{d \leq na\} = \sum_{d=0}^{[na]} \binom{n}{d} p^d (1-p)^{n-d}, \quad (5)$$

where $[na]$ denotes the largest integer less than equal to na. It can be shown that $E(f) = p$ and $E\left(\sigma_f^2\right) = p(1-p)/n$ [20]. If f is STB type, then it will have only USL (say, $USL = f_U$), and if f is LTB type, then it will have only LSL (say, $LSL = f_L$).

Suppose, m samples of size n are collected from a stable process and number of defectives observed in i^{th} sample is d_i. Then, fractions nonconforming in the i^{th} sample is $f_i = d_i/n$ $(i = 1,2,3,\ldots,m)$ and the unknown parameter p is estimated as $\hat{p} = \bar{f} = (\sum_{i=1}^{m} d_i)/mn$. So the values of $\widehat{PNL}_U$ and $\widehat{PNL}_L$can be obtained as follows:

$$\widehat{PNL}_U = P\{f > f_U\} = 1 - P\{D \leq \bar{n}f_U\} = 1 - \sum_{d=0}^{[\bar{n}f_U]} \binom{\bar{n}}{d} \bar{f}^d \left(1-\bar{f}\right)^{\bar{n}-d}, \quad (6)$$

$$\widehat{PNL}_L = P\{f < f_L\} = P\{D \leq \bar{n}f_L\} = \sum_{d=0}^{[\bar{n}f_L]} \binom{\bar{n}}{d} \bar{f}^d (1-\bar{f})^{\bar{n}-d}. \quad (7)$$

Finally, the estimates of C_u and C_l from a Binomial process are obtained as follows:

$$\hat{C}_u = (1/3) \times \Phi^{-1}\left(1 - \widehat{PNL}_U\right), \quad (8)$$

$$\hat{C}_l = (1/3) \times \Phi^{-1}\left(1 - \widehat{PNL}_L\right). \quad (9)$$

It must be noted that if $\widehat{PNU}_U$ (or $\widehat{PNL}_U$) is greater than equal to 0.5, then $\hat{C}_u$ is considered as zero. Similarly, if $\widehat{PNU}_L$ (or $\widehat{PNL}_L$) is greater than equal to 0.5, then $\hat{C}_l$ is considered as zero.

2.2. C_f Index

Yeh and Bhattacharya [16] proposed to measure C_f index as alternative to C_p. The index C_f is defined as follows:

$$C_f = \min\left(\alpha_0^L/\alpha_L, \alpha_0^U/\alpha_U\right), \tag{10}$$

where, α_0^L and α_0^U are the proportions of nonconformance the manufacturer can tolerate on the LSL and USL respectively, and α_L (=$\widehat{PNU}_L$ or $\widehat{PNL}_L$) and α_U (=$\widehat{PNU}_U$ or $\widehat{PNL}_U$) are the actual proportion of nonconformance with respect to LSL and USL respectively. In case of unilateral specification, the equivalent generalized indices for C_{pu} and C_{pl} can be represented as C_{fu} and C_{fl}, where

$$C_{fu} = \alpha_0^U/\alpha_U, \tag{11}$$

$$C_{fl} = \alpha_0^L/\alpha_L. \tag{12}$$

2.3. C_{pc} Index

Perakis and Xekalaki [17-18] proposed to measure C_{pc} index as an alternative to C_p. The index C_{pc} is defined as follows:

$$C_{pc} = (1 - p_0)/(1 - p), \tag{13}$$

where, p_0 is the minimum allowable proportion of conformance and p is the actual proportion of conformance. For unilateral specification, the equivalents generalized indices for C_{pu} and C_{pl} are C_{pcu} and C_{pcl} respectively, which can be defined as follows:

$$C_{pcu} = (1 - p_0^U)/(1 - p_U), \tag{14}$$

$$C_{pcl} = (1 - p_0^L)/(1 - p_L), \tag{15}$$

where, p_0^U and p_0^L are desired proportion of conformance with respect to USL and LSL respectively, and p_U (=$1 - \widehat{PNU}_U$ or $1 - \widehat{PNL}_U$) and p_L (=$1 - \widehat{PNU}_L$ or $1 - \widehat{PNL}_L$) are actual proportion of conformance with respect to USL and LSL respectively. It may be noted that if the actual proportion of conformance is greater than the desired proportion of conformance then C_{pcu} or C_{pcl} is greater than one and if actual proportion of conformance is less than the desired proportion of conformance then C_{pcu} or C_{pcl} is less than one. Thus, these indices can easily discriminate if a process is capable or not.

It can be observed here that $1 - p_U = \alpha_U$ and $1 - p_L = \alpha_L$. Perakis and Xekalaki [17-18] recommend that 0.9973 is a good choice for the desired proportion of conformance for both sided specifications and thus, a good choice for the desired proportion of conformance for one sided specification is 0.99865. So, $1 - p_0^U = 0.00135 = \alpha_0^U$ and $1 - p_0^L = 0.00135 = \alpha_0^L$. Thus, the indices defined by Yeh and Bhattacharya [16] and Perakis and Xekalaki [17-18] are essentially the same in case of unilateral specification. Hence, the C_{fu} (or C_{fl}) index will not be discussed further in this article.

It can be easily verified here that the index C_{pcu} or C_{pcl} will always be greater than or equal to 0.00135 and it can take very high value when the actual proportion of conformance p approaches to 1.

2.4. C_{py} Index

Maiti et al. [19] proposed to measure C_{py} index as an alternative to C_p. The index C_{py} is defined as follows:

$$C_{py} = \frac{F(USL) - F(LSL)}{1 - \alpha_0^U - \alpha_0^L} = \frac{F(U) - F(L)}{1 - \alpha_0^U - \alpha_0^L}, \quad (16)$$

where, $F(U)$ (=$1 - \widehat{PNU}_U$ or $1 - \widehat{PNL}_U$) and $F(L)$ (=$\widehat{PNU}_L$ or $\widehat{PNL}_L$) are cumulative probability distribution function of the quality characteristic at USL and LSL respectively, and α_0^U and α_0^L are the maximum allowable proportion of nonconformance at upper tail and lower tail of the distribution of the quality characteristic. Here the numerator, $F(U) - F(L)$, gives the measure of the actual process yield (i.e. actual proportion of conformance) and the denominator, $(1 - \alpha_0^U - \alpha_0^L)$ gives the measure of the desired process yield (i.e. desired proportion of conformance). It may be noted that if the actual yield is greater than the desired yield then $C_{py} > 1$ and if the actual yield is less than the desired yield then $C_{py} < 1$. Thus, one can easily judge if a process is capable or not by examining the value of C_{py} index.

Maiti et al. [19] suggested that in case of unilateral specification, median of the distribution (μ_e) should be taken as the process target and the process centre should be located such that $F(\mu_e) = [F(U) + F(L)]/2 = 1/2 = 0.5$. Therefore, for unilateral specification, generalized indices equivalent to C_{pu} and C_{pl} are expressed as

$$C_{pyu} = \frac{F(U) - F(\mu_e)}{1 - \alpha_0^U - F(\mu_e)} = \frac{F(U) - 0.5}{0.5 - \alpha_0^U}, \quad (17)$$

$$C_{pyl} = \frac{F(\mu_e) - F(L)}{F(\mu_e) - \alpha_0^L} = \frac{0.5 - F(L)}{0.5 - \alpha_0^L}, \quad (18)$$

where the value of α_0^U or α_0^L is conventionally taken as 0.00135.

It can be easily verified here that the index C_{pyu} or C_{pyl} can take a maximum value of 1.0027. Also, if $F(U)$ is less than or equal to 0.5, then C_{pyu} is considered as zero. Similarly, if $F(L)$ is greater than or equal to 0.5, then C_{pyl} is considered as zero.

3. Assessing Relative Goodness of the Estimates of Different Generalized Indices

It is always possible to derive the relation between the estimate of an index obtained from a process and the Expected Proportion of Conformance (EPC) in the process outputs. As mentioned earlier, Poisson or binomial data usually have one-sided specification and therefore, here four different process capability indices with respect to USL only (e.g.C_{pu}, C_u, C_{pcu}, C_{pyu}) are considered for derivation of these relationships. Among these, C_{pu} is the conventional process capability index with respect to USL and C_u, C_{pcu} and C_{pyu} are the generalized process capability indices with respect to USL. The EPC in the process outputs corresponding to an estimate of C_{pu} may be represented as EPC-C_{pu}. Similarly, the EPC in the process outputs corresponding to the estimates of C_u, C_{pcu} and C_{pyu} may be represented as EPC-C_u, EPC-C_{pcu} and EPC-C_{pyu} respectively. The derived relationships between EPCs in the process outputs and the estimates of these indices are shown in ***Eqs. (19)-(22)***.

$$\text{EPC } C_{pu}\text{-} = \Phi\left(3 \times \hat{C}_{pu}\right). \qquad \textbf{\textit{(19)}}$$

$$\text{EPC-}C_u = \Phi\left(3 \times \hat{C}_u\right). \qquad \textbf{\textit{(20)}}$$

$$\text{EPC-}C_{pcu} = 1 - 0.00135/\hat{C}_{pcu}\ ;\ \hat{C}_{pcu} \geq 0.00135. \qquad \textbf{\textit{(21)}}$$

$$\text{EPC-}C_{pyu} = 0.5 + \hat{C}_{pyu} \times 0.49865; \hat{C}_{pyu} \leq 1.0027. \qquad \textbf{\textit{(22)}}$$

It may be noted that if the estimate of an index (conventional or generalized) is equal to one, it always indicates that EPC in the process outputs is 0.99865. However, if the estimates of the indices are different from one, the same value of different indices signifies different EPC in the process outputs. For example, a value of 0.5 for C_{pu}, C_u, C_{pcu} and C_{pyu} indicates that EPC in the process outputs are 0.93319, 0.93319, 0.99730 and 0.74933 respectively. Conversely, the estimates of different indices obtained from the same process will be different. For example, if the expected proportion of conformance in a process is 0.99730 and C_{pu}, C_u, C_{pcu} and C_{pyu} are evaluated from the same process, the values of $\hat{C}_{pu}$, $\hat{C}_u$, $\hat{C}_{pcu}$ and $\hat{C}_{pyu}$ will be about 0.9274, 0.9274, 0.50 and 0.9973 respectively. Similarly, if the expected proportion of conformance in a process is 0.95, the values of $\hat{C}_{pu}$, $\hat{C}_u$, $\hat{C}_{pcu}$ and $\hat{C}_{pyu}$ will be about 0.5483, 0.5483, 0.027 and 0.9024 respectively.

Over the years, process managers, engineers and other decision makers have become accustomed to evaluation of process capability indices from normal processes and its interpretation. Due to legacy of usages of conventional process capability indices and its interpretations, a user of a generalized index may tend to interpret its values with reference to the values of the conventional process capability indices for the bad, good or very good capable normal processes, and thus he/she may get a false impression about the capability of the process by examining the estimate of a generalized index. Therefore, goodness of a generalized index should be judged from the perspective of potential chances of getting false impression about the capability of a process based on the estimate of the index. Ideally, the same values of the conventional index (C_{pu}) and different generalized indices (C_u, C_{pcu}, and C_{pyu}) should lead to similar interpretation about the capability of a process, and the generalized index whose values are closer to the values of conventional index (C_{pu}) at different conformance levels to the specification should be considered better in quantifying process capability.

For understanding the relative goodness of different generalized indices, EPCs corresponding to the same value of conventional index (C_{pu}) and generalized indices (C_u, C_{pcu}, and C_{pyu}) are computed using ***Eqs. (19)-(22)***. These computed EPCs for different index values spanning over 0.1 to 1.0 are presented in ***Table 1***. In this Table, the value in the second cell of a row is the computed EPCs when C_{pu} = value in first cell of that row. Similarly, the value in the third cell of a row is the computed EPCs when C_u = value in first cell of that row, etc.

Table 1. EPCs corresponding to the same value of different indices.

Index value	EPCs corresponding to the same value of different indices			
	EPC-C_{pu}	EPC-C_u	EPC-C_{pcu}	EPC-C_{pyu}
0.01	0.51197	0.51197	0.86500	0.50499
0.02	0.52392	0.52392	0.93250	0.50997
0.03	0.53586	0.53586	0.95500	0.51496
0.04	0.54776	0.54776	0.96625	0.51995
0.05	0.55962	0.55962	0.97300	0.52493
0.06	0.57142	0.57142	0.97750	0.52992
0.07	0.58317	0.58317	0.98071	0.53491
0.08	0.59483	0.59483	0.98313	0.53989
0.09	0.60642	0.60642	0.98500	0.54488
0.1	0.61791	0.61791	0.98650	0.54987
0.2	0.72575	0.72575	0.99325	0.59973
0.3	0.81594	0.81594	0.99550	0.64960
0.4	0.88493	0.88493	0.99663	0.69946
0.5	0.93319	0.93319	0.99730	0.74933
0.6	0.96407	0.96407	0.99775	0.79919
0.7	0.98214	0.98214	0.99807	0.84906
0.8	0.9918	0.9918	0.99831	0.89892
0.9	0.99653	0.99653	0.99850	0.94879
1.0	0.99865	0.99865	0.99865	0.99865

Figure 1 shows the graphs of EPC in the process outputs versus index value for the conventional index C_{pu} and the generalized indices C_u, C_{pcu} and C_{pyu}. It can be noted from ***Table 1*** and ***Figure 1*** that when the index value is equal to one the conventional index (C_{pu}) as well as all the generalized indices indicate the same EPC in the process outputs, i.e. a value of one for the conventional index (C_{pu}) and the generalized indices (C_u, C_{pcu} and C_{pyu}) bear the same interpretation. It is further noted that interpretation of values of the conventional index (C_{pu}) and the generalized index (C_u) match exactly well throughout the considered range of index values. However, interpretations of the values of C_{pcu} and C_{pyu} tend to differ from the values of C_{pu} and C_u as the index values become more and more less than one.

The graphs of EPC versus index value for C_{pu} and C_u are noted to be coincided over one another, which imply that interpretations of different values of C_{pu} and C_u are the same. This graph is found to be a nonlinear curve. Slope of the curve is high for lower index values (up to 0.5) and the slope approaches zero as the index value approaches one. This implies that for a small increase in index value, the increase in EPC is much more initially but as the index value approaches one, for a large increase in index value the increase in EPC is quite small. On the other hand, the graph of EPC versus index value for C_{pyu} is found to be linear. This implies that EPC increases uniformly throughout the range of index values. However, the graph of EPC versus index value for the C_{pcu} is observed to be the most peculiar and unacceptable. Even a very small value of C_{pcu} signifies a very high value of EPC. For example, $C_{pcu} = 0.01$ signifies EPC = 0.86500. This curve increases very steeply for index values from 0.01 to 0.06 and then flattens and slope becomes almost zero. Thus, for a large increase in index value the increase in EPC is quite small. In other words, a wide range of values of C_{pcu} will indicate almost the same EPC in a process, which is undesirable.

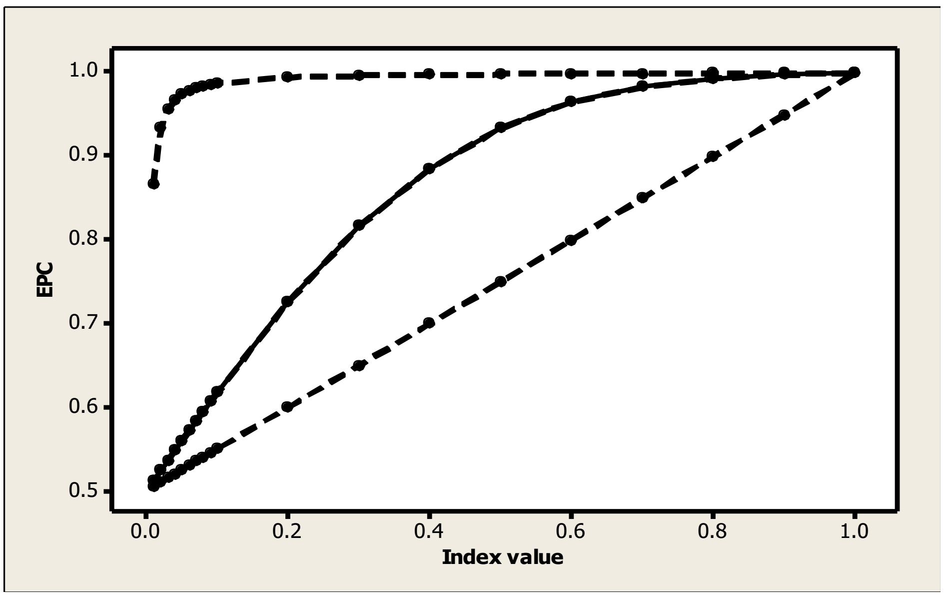


Figure 1. EPC in the process outputs indicated by different values of the indices.

Now let us consider that the proportion of conformance in a process is 0.9 and the generalized indices C_u, C_{pcu} and C_{pyu} are estimated from this process. The approximate values of these estimates can be obtained by drawing a straight line parallel to X-axis at EPC = 0.9 in ***Figure 1***. It can be observed from ***Figure 1*** that the values of C_u, C_{pcu} and C_{pyu} will be about 0.427, 0.0135 and 0.802 respectively. Obviously, if one examines the process capability index value as C_{pcu}=0.0135, he/she will get an impression that the capability of the process is very poor. Again, if one examines the process capability index value as C_{pyu}= 0.802, he/she will get an impression about the capability of the same process that it is not so bad. On the other hand, it can be observed from ***Figure 1*** that for the same index value, the EPC corresponding to C_{pcu} is always more than the EPC corresponding to C_u and the EPC corresponding to C_{pyu} is always less than the EPC corresponding to C_u. In fact, the generic process capability index C_{pcu} underestimate the capability of a process and the generic process capability index C_{pyu} overestimate the capability of a process. Only the generic index C_u properly quantifies the capability of a process.

4. Analysis and Related Results

Analysis of three sets of case study data taken from Poisson processes and two sets of case study data taken from binomial processes are presented in this section. In all these case studies only USLs are specified.

4.1. Analysis of Poisson Data

4.1.1. Case study 1

Maravelakis [21] considered a manufacturing process for illustrating transformation approach for process capability index. In this case study, 100 units of products were collected, number of occurrences of defects on each product were observed and recorded. The USL for the number of defects in a unit was $c_U = 9$. In the data set, the total number of defects was found to be $\sum_{i=1}^{100} c_i = 506$. So, unknown parameter λ was estimated as $\hat{\lambda} = \bar{c} = 506/100 = 5.06$. The defects data c_i $(i = 1,2,3, \ldots, 100)$ were plotted in a c-chart, which revealed that during sample collection the process was in control.

These data are analyzed first for obtaining the estimate of C_u. Here, $\hat{\lambda} = \bar{c} = 5.06$ and $c_U = 9$. So PNU_U and Z_U values are obtained as follows:

$$PNU_U = 1 - \sum_{c=0}^{9} e^{-5.06}(5.06)^c/c! = 1 - 0.96594 = 0.03406,$$

$$Z_U = \Phi^{-1}(1 - PNU_U) = \Phi^{-1}(1 - 0.03406) = \Phi^{-1}(0.96594) = 1.824.$$

Thus, C_u is estimated as

$$\hat{C}_u = (1/3) \times Z_U = (1/3) \times 1.824 = 0.6081.$$

This implies that the EPC in the concerned Poisson process is $\Phi(3 \times \hat{C}_u) = \Phi(3 \times 0.608) \cong \Phi(1.82) = 0.96562$.

Now C_{pcu} and C_{pyu} are estimated from the same data. The desired proportion of conforming units with respect to USL, i.e. p_0^U is not specified and so as per convention, it is taken as 0.99865. This implies that $(1 - p_0^U) = 0.00135$. The actual proportion of nonconforming units, $(1 - p_U)$ is computed to be $\widehat{PNU}_U = 0.03406$. Therefore, the process capability index C_{pcu} is estimated as

$$\hat{C}_{pcu} = 0.00135/0.03406 = 0.0396.$$

The estimated $\hat{C}_{pcu}$ value is less than 0.05 and thus, it gives an impression that the process capability is very poor. However, in reality, about 96.56% of outputs in this process conform to the specification.

On the other hand, as per convention, the maximum allowable proportion of nonconforming units at upper tail, α_0^U is considered to be 0.00135. The cumulative probability up to the $USL = c_U = 9$, $F(U)$ is computed as 0.96594. Therefore, C_{pyu} is estimated as

$$\hat{C}_{pyu} = (0.96594 - 0.5)/(0.5 - 0.00135) = 0.46594/0.49865 = 0.9344.$$

The estimated $\hat{C}_{pyu}$ value is quite close to 1.0 and thus, it gives an impression that the capability of the process is not very bad. However, in reality, about 3.44% of process outputs do not conform to the specification.

4.1.2. Case study 2

Montgomery [20] presented a set of process data on number of defects observed on 20 inspection units of Printed Circuit Boards (PCBs) for illustration of attribute process control charts. Each inspection unit consisted of successive samples of five PCBs and each inspection unit was collected after every one hour. The total number of defects was found to be 160 and therefore, the estimate of the unknown parameter was obtained as $\hat{\lambda} = \bar{c} = 160/20 = 8$. The c-chart of the defects data indicates that the process was in statistical control during sample collection. So process capability indices are estimated from these data.

Montgomery [20] did not specify USL for the number of defects in a PCB. For the purpose of estimation of process capability indices, we assume that $\bar{c} + 2 \times \sqrt{c} = 8 + 5.66 \approx 14$ is the USL for the number of defects in a PCB, i.e. $c_U = 14$. At first, the estimate of C_u are obtained from these data. It is found that $\hat{C}_u = 0.7047$, which implies that the EPC in this Poisson process is $\Phi(3 \times \hat{C}_u) = \Phi(3 \times 0.7047) \cong \Phi(2.11) = 0.98257$.

Now C_{pcu} and C_{pyu} are estimated from the same data. It is found that $\hat{C}_{pcu} = 0.0782$, which is less than 0.1 and thus, gives an impression that the process capability is very poor although about 98.26% of process outputs conform to the specification. On the other hand, it is found that $\hat{C}_{pyu} = 0.9681$, which is close to 1.0 and thus, it gives an impression that the capability of the process is quite good. However, in reality, about 1.74% of process outputs do not conform to the specification.

4.1.3. Case study 3

The inspection results on 25 wafers each was containing 100 chips were presented in the NIST/SEMATECH e-handbook of statistical methods [22] for illustration of construction of c-chart. The total number of defects was found to be 400 and the estimate of the unknown parameter was obtained as $\hat{\lambda} = \bar{c} = 400/25 = 16$. The c-chart of the defects data indicated that the process was in statistical control during sample collection. So process capability indices are now estimated from these data.

Here, USL for the number of defects in a wafer is not specified. So, we assume that $\bar{c} + 2 \times \sqrt{\bar{c}} = 16 + 2\sqrt{16} = 24$ is the USL for the number of defects in a wafer, i.e. $c_U = 24$. At first the estimate of C_u is obtained. It is found that $\hat{C}_u = 0.6694$, which implies that EPC in the Poisson process is $\Phi(3 \times \hat{C}_u) = \Phi(3 \times 0.6694) \cong \Phi(2.01) = 0.97778$.

Now C_{pcu} and C_{pyu} are estimated from the same data. It is found that $\hat{C}_{pcu} = 0.0601$, which is less than 0.1 and thus, it gives an impression that the process capability is very poor although in reality about 97.78% of process outputs conform to the specification. On the other hand, it is found that $\hat{C}_{pyu} = 0.9580$, which is quite close to 1.0 and thus, it gives an impression that the capability of the process is not very bad. However, in reality, about 2.22% of process outputs do not conform to the specification.

The results of analysis of the three case studies are summarized in ***Table 2***. The results in ***Table 2*** shows that the estimated $\hat{C}_{pcu}$ values in the three case studies vary from 0.0396 to 0.0782, which are less than 0.1 and thus, give an impression that capabilities of these processes is very poor. However, in reality, expected proportions of conformance in these processes vary from about 96.56% to 98.26%, which are not so bad. This implies that the generalized index C_{pcu} usually underestimate the capability a Poisson process. On the other hand, the estimated $\hat{C}_{pyu}$ values in the three case studies vary from 0.9344 to 0.9681, which are close to 1.0 and thus give an impression that capabilities of these processes are quite well. However, in reality, expected percentages of nonconformance in these processes vary from about 1.74% to 3.44%, which are quite high. This implies that the generalized index C_{pyu} usually overestimate the capability of a Poisson process.

Table 2. Estimates of generalized indices and the impression about capability of the process indicated by these estimates.

Case Study	Generalized index	Estimated value	Impression about the capability of the process
1	C_u	0.6081	EPC in the process outputs is about 96.56%
	C_{pcu}	0.0396	Process capability is very poor
	C_{pyu}	0.9344	Process capability is quite good
	C_u	0.7047	EPC in the process outputs is about 98.26%
2	C_{pcu}	0.0782	Process capability is very poor
	C_{pyu}	0.9681	Process capability is quite good
	C_u	0.6694	EPC in the process outputs is about 97.78%
3	C_{pcu}	0.0601	Process capability is very poor
	C_{pyu}	0.9580	Process capability is quite good

4.2. Analysis of Binomial Data

4.2.1. Case study 1

Maravelakis [21] considered a manufacturing process for illustrating transformation technique for binomial data and subsequent computation of process capability index. Maravelakis [21] collected a total of $m = 100$ samples each of size $n = 30$ from the manufacturing process, and observed the number of nonconforming items (d) in each sample. The fraction (proportion) nonconforming in these samples are calculated and plotted in p-chart. The plotted p-chart revealed that the process was in control. So he used the data for illustrating the transformation approach for measuring process capability index from binomial data. The USL for the fraction nonconforming is $f_U = 0.2$. The total number of nonconforming items in these samples is found to be $\sum_{i=1}^{m} d_i = 286$ and so, the average fraction nonconforming is computed as $\bar{f} = 286/(30 \times 100) = 0.09533$. All the generalized process capability indices are now estimated from these data.

At first the estimate of C_u is obtained. Here, $n = 30$, $\bar{f} = 0.09533$, $f_U = 0.2$ and $[nf_U] = 6$. Therefore, PNL_U and Z_U are computed as follows:

$$PNL_U = 1 - \sum_{d=0}^{6} \binom{30}{d} (0.09533)^d (0.90467)^{30-d} = 0.02039,$$

$$Z_U = \Phi^{-1}(1 - PNL_U) = \Phi^{-1}(1 - 0.02039) = \Phi^{-1}(0.97961) \cong 2.05.$$

Thus, C_u is estimated as

$$\hat{C}_u = (1/3) \times Z_U = (1/3) \times 2.05 = 0.6833.$$

This implies that the expected proportion of conformance in the concerned binomial process is $\Phi\left(3 \times \hat{C}_u\right) = \Phi(3 \times 0.6833) \cong \Phi(2.05) = 0.97982$

Now C_{pcu} and C_{pyu} are estimated from the same data. The desired proportion of conforming lots with respect to USL, i.e. p_0^U is not specified and so as per convention, it is taken as 0.99865. This implies that acceptable proportion of nonconforming lots, $\left(1 - p_0^U\right) = 0.00135$. Here the actual proportion of nonconforming lots, $(1 - p_U)$ is computed to be $\widehat{PNL}_U = 0.02039$. Therefore, the process capability index C_{pcu} is estimated as

$$\hat{C}_{pcu} = 0.00135/0.02039 = 0.0662.$$

The estimated $\hat{C}_{pcu}$ value is less than 0.1 and thus, it gives an impression that the process capability is very poor. However, in reality, about 97.98% of process outputs conform to the specification.

On the other hand, as per convention, the maximum allowable proportion of nonconforming lots at upper tail, i.e. α_0^U is considered to be 0.00135. The cumulative probability of conforming lots up to the USL, i.e. $F(U)$ is computed as 0.97961. Therefore, C_{pyu} is estimated as

$$\hat{C}_{pyu} = (0.97961 - 0.5)/(0.5 - 0.00135) = 0.47961/0.49865 = 0.9618.$$

The estimated $\hat{C}_{pyu}$ value is close to 1.0 and thus, it gives an impression that the capability of the process is quite well. However, in reality, 2.02% of process outputs do not conform to the specification.

4.2.2. Case study 2

Montgomery [20] presented, in exercise 7.3 (pp. 335), a set of process data on total number of personal computers inspected and total number of nonconforming personal computer observed in each day over last ten consecutive days. Then he wanted to know if the process was in control. The plotted fraction nonconforming control chart exhibited that the process was in control, and therefore, it is decided to use the same data for process capability analysis purpose.

In this data set, sample size (n_i) was variable and the average sample size ($\bar{n}$) is found to be 100, and the average fraction nonconforming ($\bar{f}$) is found to be 0.06. Montgomery [20] did not specify the USL for the fraction nonconforming. However, USL needs to be known for carrying out process capability analysis. For the purpose of process capability analysis, here we assume that $\bar{f} + 2 \times \sqrt{\bar{f}(1-\bar{f})/\bar{n}} \approx 0.10$ is the USL for the fraction nonconforming, i.e.$f_U = 0.10$.

The estimate of C_u is obtained first. It is found that $\hat{C}_u = 0.5931$, which implies that the expected proportion of conformance in the concerned binomial process is $\Phi(3 \times \hat{C}_u) = \Phi(3 \times 0.5931) \cong \Phi(1.78) = 0.96246$.

Now C_{pcu} and C_{pyu} are estimated from the same data. It is found that $\hat{C}_{pcu} = 0.0360$, which is less than 0.05 and thus, gives an impression that the process capability is very poor although about 96.25% of process outputs conform to the specification. On the other hand, it is found that $\hat{C}_{pyu} = 0.9273$, which is close to 1.0 and thus, it gives an impression that the capability of the process is not so bad. However, in reality, 3.75% of process outputs do not conform to the specification.

The results of analysis of the two case studies are summarized in ***Table 3***. The results in ***Table 3*** shows that the estimated $\hat{C}_{pcu}$ values in the two case studies vary from 0.0360 to 0.0660, which give an impression that capabilities of these processes is very poor. However, in reality, expected proportions of conformance in these processes vary from 0.9625 to 0.9798, which are not so bad. On the other hand, the estimated $\hat{C}_{pyu}$ values in these case studies vary from 0.9273 to 0.9620, which are close to 1.0 and thus give an impression that capabilities of these processes are quite well. However, in reality, expected percentages of nonconformance in these processes vary from about 2.02% to 3.75%, which are quite high. This implies that the generalized index C_{pcu} usually underestimate the capability a binomial process whereas the generalized index C_{pyu} usually overestimate the capability of a binomial process.

Table 3. Estimates of generalized indices and the impression about capability of the process indicated by these estimates.

Case study	Generalized index	Estimated value	Impression about the capability of the process
1	C_u	0.6820	EPC in the process outputs is about 0.9798
	C_{pcu}	0.0662	Process capability is very poor
	C_{pyu}	0.9618	Process capability is quite good
2	C_u	0.5931	EPC in the process outputs is about 0.9625
	C_{pcu}	0.0360	Process capability is very poor
	C_{pyu}	0.9273	Process capability is quite good

5. Discussions

It has been observed that curves for C_{pu} and C_u shown in ***Figure 1*** coincide over one another. This implies that same values of $\hat{C}_u$ (or $\hat{C}_l$) from a Poisson/binomial process and $\hat{C}_{pu}$ (or $\hat{C}_{pl}$) from a normal process bear the same interpretation. This happens because of the following facts. The $\hat{C}_u$ value in a Poisson (or binomial) process is computed by directly mapping the fraction of nonconforming units with respect to USL (or the fraction of nonconforming lots with respect to USL) to the *Z*-value (say, Z_U) of standard normal distribution that results in the same probability of nonconformance in the upper tail and one-third of this Z_U-value is considered as an estimate of C_u. On the other hand, the $\hat{C}_{pu}$ value in a normal process is computed as $\hat{C}_{pu} = (USL - \hat{\mu})/3\hat{\sigma} = \frac{1}{3}Z_U$. Thus, same values of $\hat{C}_u$ (or $\hat{C}_l$) from a Poisson/binomial process and $\hat{C}_{pu}$ (or $\hat{C}_{pl}$) from a normal process always bear the same interpretation.

The problem with the C_{pcu} (or C_{pcl}) index is that it is estimated as ratio of two very small numbers, where numerator is 0.00135 (acceptable proportion of nonconformance) and denominator is actual proportion of nonconforming units/lots with respect to USL (or LSL). Thus, the estimate is highly impacted due to a minor deviation in the value of actual proportion of nonconformance from the acceptable proportion of nonconformance. For example, if actual proportion of nonconformance is 0.00135 then the value of $\hat{C}_{pcu}$ is equal to one but if the actual proportion of nonconformance becomes 0.0001 then the value of $\hat{C}_{pcu}$ would become as high as 13.5, which would give a misleading impression that the process is highly capable. On the other hand, if the actual proportion of nonconformance becomes 0.005 then the value of $\hat{C}_{pcu}$ would become as low as 0.27, which again gives a misleading impression that the process capability is very poor.

The C_{pyu} (or C_{pyl}) index suffers from another problem. The values of the ratios $[F(U) - 0.5]/(0.5 - \alpha_0^U)$ and $[0.5 - F(L)]/(0.5 - \alpha_0^L)$ are considered as the estimate of C_{pyu} and C_{pyl} respectively. Since the values α_0^U and α_0^L are usually taken as 0.00135, the denominator is always equal to 0.49865 in both the ratios. On the other hand, the values of the numerators in both the ratios can be at most 0.5. Therefore, the maximum value of $\hat{C}_{pyu}$ (or $\hat{C}_{pyl}$) in a process can be $0.5/0.49865 = 1.0027$. This implies that the estimate of the generalized index C_{pyu} (or C_{pyl}) would fail to make distinction between just capable process and highly capable process.

6. Conclusions

Process Capability Analysis (PCA) is an important analytic tool that is used to assess if a process is capable of meeting the specified requirements. In most applications of PCA, the quality characteristic is assumed to follow normal distribution. But, in practice, many quality characteristics follow Poisson or binomial distributions, and these characteristics usually have one-sided specification. In such cases, computations of C_{pu} or C_{pl} using the standard formula is inappropriate. In order to alleviate the problem, some generalized indices (e.g. C_u and C_l, C_{pcu} and C_{pcl}, C_{pyu} and C_{pyl}**)** are proposed in literature. In this paper, relative goodness of these indices in quantifying the capability of a process is assessed. It is found that only $\hat{C}_u$ (or $\hat{C}_l$) gives proper assessment about the capability of a process. All other generalized indices give a false impression about the capability of a process and thus usages of those indices should always be avoided. The results of analysis of multiple case study data taken from Poisson and binomial processes validate the above findings.

Acknowledgement

We are sincerely grateful to the referees for their valuable comments which have helped in improving the overall clarity and readability of this article.

References

[1] Ryan, T. P. (2011). *Statistical methods for quality improvement*. John Wiley & Sons.
[2] Kane, V. E. (1986). Process capability indices. *Journal of quality technology*, *18*(1), 41-52.
[3] Kotz, S., & Johnson, N. L. (1993). *Process capability indices*. CRC press.
[4] English, J. R., & Taylor, G. D. (1993). Process capability analysis—a robustness study. *The international journal of production research*, *31*(7), 1621-1635.
[5] Kotz, S., & Johnson, N. L. (2002). Process capability indices—a review, 1992–2000. *Journal of quality technology*, *34*(1), 2-19.
[6] Chen, K. S., Huang, M. L., & Hung, Y. (2008). Process capability analysis chart with the application of Cpm. *International journal of production research*, *46*(16), 4483-4499.
[7] Wu, C. W., Pearn, W. L., & Kotz, S. (2009). An overview of theory and practice on process capability indices for quality assurance. *International journal of production economics*, *117*(2), 338-359.
[8] Yum, B. J., & Kim, K. W. (2011). A bibliography of the literature on process capability indices: 2000–2009. *Quality and reliability engineering international*, *27*(3), 251-268.
[9] Chen, K. S., Wang, K. J., & Chang, T. C. (2017). A novel approach to deriving the lower confidence limit of indices C pu, C pl, and C pk in assessing process capability. *International journal of production research*, *55*(17), 4963-4981.
[10] Clements, J. A. (1989). Process capability calculations, for non-normal distributions. *Quality progress*, *22*(9), 95-100.

[11] Pearn, W. L., & Chen, K. S. (1995). Estimating process capability indices for non-normal pearsonian populations. *Quality and reliability engineering international*, *11*(5), 386-388.
[12] Shore, H. (1998). A new approach to analysing non-normal quality data with application to process capability analysis. *International journal of production research*, *36*(7), 1917-1933.
[13] Chen, J. P. (2000). Re-evaluating the process capability indices for non-normal distributions. *International journal of production research*, *38*(6), 1311-1324.
[14] Goswami, A., & Dutta, H. N. (2013). Some studies on normal and non-normal process capability indices. *Int. j. math. stat. invent*, *1*(2), 31-40.
[15] Borges, W. D. S., & Ho, L. L. (2001). A fraction defective based capability index. *Quality and reliability engineering international*, *17*(6), 447-458.
[16] Yeh, A. B., & Bhattcharya, S. (1998). A robust process capability index. *Communications in statistics-simulation and computation*, *27*(2), 565-589.
[17] Perakis, M., & Xekalaki, E. (2002). A process capability index that is based on the proportion of conformance. *Journal of statistical computation and simulation*, *72*(9), 707-718.
[18] Perakis, M., & Xekalaki, E. (2005). A process capability index for discrete processes. *Journal of statistical computation and simulation*, *75*(3), 175-187.
[19] Maiti, S.S., Saha, M., & Nanda, A.K. (2010). On generalising process capability indices. *Journal of quality technology and quantitative management*, 7(3), 279-300.
[20] Montgomery, D.C. (2009). *Introduction to statistical process control*, *6th edition*. Jefferson City, USA: John Wiley & Sons, Inc.
[21] Maravelakis, P. E. (2016). Process capability indices for data following the Poisson or binomial distribution. *Quality technology & quantitative management*, *13*(2), 197-206.
[22] Sematech, N. I. S. T. (2012). *NIST/SEMATECH e-handbook of statistical methods*, http://www.itl.nist.gov/div898/handbook/