IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 19, 2016, accepted October 10, 2016, date of publication October 24, 2016, date of current version November 18, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2620241

Capability Indices for Non-Normal Distribution Using Gini's Mean Difference as Measure of Variability

MUHAMMAD KASHIF[1], MUHAMMAD ASLAM[1], ALI HUSSEIN AL-MARSHADI[1], AND CHI-HYUCK JUN[2]
[1]Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia
[2]Department of Industrial and Management Engineering, Pohang University of Science and Technology, Pohang 790-784, South Korea

Corresponding author: M. Aslam (aslam_ravian@hotmail.com)

This work was supported by the Deanship of Scientific Research, King Abdulaziz University, Jeddah.

ABSTRACT This paper investigates the efficiency of Gini's mean difference (GMD) as a measure of variability in two commonly used process capability indices (PCIs), i.e., Cp and Cpk. A comparison has been carried out to evaluate the performance of GMD-based PCIs and Pearn and Chen quantile-based PCIs under low, moderate, and high asymmetry using Weibull distribution. The simulation results, under low and moderate asymmetric condition, indicate that GMD-based PCIs are more close to target values than quantile approach. Beside point estimation, nonparametric bootstrap confidence intervals, such as standard, percentile, and bias corrected percentile with their coverage probabilities also have been calculated. Using quantile approach, bias corrected percentile (BCPB) method is more effective for both Cp and Cpk, where as in case of GMD, both BCPB and percentile bootstrap method can be used to estimate the confidence interval of Cp and Cpk, respectively.

INDEX TERMS Gini's mean difference, process capability indices, non-normal, Weibull distribution.

I. INTRODUCTION

Process capability index is a major tool to evaluate the manufacturing progress of any process. The traditional PCIs such as, C_p ,C_{pk}, C_{pm} and C_{pmk} performed well when process follows the normal behavior [1], [2]. In addition, non-normal distribution process is also being practiced in an industrial environment. Therefore, both normal and non-normal processes capability indices are frequently used to monitor the process performance.

A. NORMAL PROCESS CAPABILITY INDEX

The most commonly used PCIs are the C_p given by Juran [3] and the C_{pk}, given by Kane [4]. The index C_p, which is related to the upper and lower specification limits, is defined as

$$C_p = \frac{USL - LSL}{6\sigma} \tag{1}$$

On the other hand, the index C_{pk}, which is more sensitive to departures from normality than C_p, for normal behavior is given by

$$C_{pk} = min\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right) \tag{2}$$

In (2), μ is the population mean and σ is the population standard deviation, which is estimated by the sample standard deviation when it is unknown. The standard deviation basically represents the process variability which may be short term or long term. The commonly used PCIs are based on both short term and long term variability. The indices C_p and C_{pk} are referred as short term PCIs, whereas the indices P_p and P_{pk} are considered as long term PPIs [5].

The PCIs used in industry provide a single numerical measure which indicates the process performance. If the resulting value of PCI (either C_p or C_{pk}) is < 1.00, the process is called inadequate. The process is called capable if $1.00 \leq PCI \leq 1.33$. For satisfactory it should be $1.33 \leq PCI \leq 1.50$, and considered super if $PCI \geq 2$ [6].

In application point of view, different experts recommended different values for existing and new processes.

In general, C_p value equal to 1.33 is recommended for existing processes and C_p value 1.67 or higher for new processes. Some authors, considered, $C_p = 1.33$ for existing processes, $C_p = 1.50$ for new processes and $C_p = 1.67$ for safety, critical parameters and new processes for two-sided specifications [7]. In present study, the range of PCI is considered 1.67 or higher.

B. NON NORMAL PROCESS CAPABILITY INDEX

However, due to different noisy, complex and multifunctional behavior of any factor, many processes in practice are non-normal [2]. The non-normality effects the efficiency of both sample mean and standard deviation and they are not considered as meaningful estimators to deal with such situation. Therefore, the PCIs defined in eq.1 and eq.2 would not be reliable and may give erroneous and misleading results. Thereby, it is necessary to take into account the non-normality to prevent the loss of resources, money and time, hence practitioners made an accurate result [8].

To deal with non-normal processes, many researchers focused on different methodologies [7]. The reliable estimators of non-normal PCIs are obtained by using two approaches. The first one is to transform the non-normal data into normal for the use of normal based PCIs. The second approach is to use PCIs defined for non-normal data [6].

In transformation methods, Box-Cox power transformation, Johnson transformation system and Clements methods using pearson curves are used. On the other hand, empirical distribution method, modification of existing PCIs and alternative measures of variability are commonly used methods for second approach [6], [7]. There are many studies in which researchers have made comparisons with in each approach or compared both approaches at a time for dealing with non-normal PCIs [2], [5], [6], [8], [9]. All these methods have been criticized by the researchers because of their variable performance under different situations. So, no single method has been recommended that works accurately in all situations [2].

Senvar and Kahraman [9] proposed the percentile based basic PCIs for non-normal data and then developed fuzzy formulation using Clements method. The performance of proposed PCIs are compared using Weibull distribution. Later on in another study, Senvar and Kahraman [8] introduced type-2 fuzzy percentile based PCIs for non-normal data via Clements methods and then compared with their crisp types. The comparison showed that proposed PCIs are more informative, sensitive and flexible to evaluate the process performance.

Sennaroglu and Senvar [6] presents a comparison of Box-Cox transformation and weighted variance methods for non-normal process capability index using Weibull distribution. Based on various summary statistics, they concluded that Box-Cox transformation method produces better estimates for process capability index than weighted variance method.

Recently, Senvar and Sennaroglu [5] compared Clements approach, Box-cox transformation and Johnson transformation method for handling non-normal PCI when data follow Weibull distribution. The Weibull distributed data with different parameters are used to figure out the effect of the tail behaviors on PCIs. Based on different measures like box-plot, descriptive statistics, the root mean square deviation and a radar chart, they concluded that Clements approach is the best among three methods.

The transformation approach has ability to produce good results as pointed out by [10] but it does not become very popular among practitioners because of extensive computing and translating the computed results with regards to the original scales [5]. In this regard, Clements [11] introduced the concept of quantile using person family of distribution for estimating the standard PCIs. Due to simplicity in calculation and application, this approach is one of the most popular one for dealing with non-normality [8].

For non-normal distribution, the PCIs defined in (1) and (2) should be modified [5]. A widely adopted procedure to construct non-normal PCIs is to substitute 6σ in (1) by the range $R = \text{Ʉ} - \text{Ƚ}$ which covers 99.73% of the distribution of the monitored process data, where Ʉ and Ƚ are the 0.135^{th} and 99.865^{th} quantiles of the corresponding non-normal distribution, respectively. This idea is introduced by Clements [11] and further modified by Pearn and Chen [12], who replaced 3σ in (2) by $[\text{Ʉ} - \text{Ƚ}/2]$. Based on modified approach [12], the index C_p and C_{pk} can be defined as

$$C^*_{Np} = \frac{USL - LSL}{\text{Ʉ} - \text{Ƚ}} \tag{3}$$

$$C^*_{Npk} = 2 * min\left(\frac{USL - \text{ɱ}}{\text{Ʉ} - \text{Ƚ}}, \frac{\text{ɱ} - LSL}{\text{Ʉ} - \text{Ƚ}}\right) \tag{4}$$

where ɱ is the 50.00^{th} quantiles of the corresponding non-normal distribution. In the modified C_p and C_{pk} defined in (3) and (4), the center of the process is based on the median, because median is a robust measure of the central tendency than mean particularly for skewed distributions.

C. ROBUST MEASURES OF VARIABILITY

But it is well established that the use of these PCIs, for heavily skewed distributions, did not provide accurate results [13]–[15]. So in this case, several authors in literature, have promoted the use of other robust measures of variability such as median absolute deviation, interquartile range and Gini's mean difference [15], [16]. Among these robust measures, GMD is considered as a universal estimator of standard deviation due to its less sensitivity to outliers, but its extensive application as a measure of variability has been rendered because of few arising computing issues i.e. estimating the variance of its estimator [16]–[18]. The GMD was developed by Professor Carrodo Gini [19] for measuring variability of the non-normal data. Later on, many authors [16], [20], [21] showed that GMD was more informative and effective measure of variability than standard deviation for highly skewed data. Therefore, the fundamental objectives of this study are, (i) to use the Gini's mean difference as a measure of variability in two commonly used PCIs C_p and C_{pk}.

(ii) to compare the performance of modified PCIs with existing quantile based PCIs, and (iii) to examine how asymmetric levels of the distribution along with sample size affect the accuracy of these PCIs.

D. PAPER ORGANIZATION

The rest of the study is organized as follows. Section 2 clearly demonstrates the procedure of the GMD based process capability indices. Section 3 and 4 will employ the simulation study and numerical example to demonstrate the effectiveness of the proposed approach. Concluding remarks are finally made in section 4.

E. ABBREVIATIONS AND ACRONYMS

GMD	Gini's Mean Difference
PCI	Process Capabilities Indices
PC	Pear and Chen Quantile Method
USL	Upper Specification Limit
LSL	Lower Specification Limit
σ	Population Standard Deviation
μ	Population Mean
ɱ	50^{th} Quantile of the Corresponding Distribution
Ƚ	0.00135^{th} Quantile of the Corresponding Distribution
Ʉ	0.99865^{th} Quantile of the Corresponding Distribution
CI	Confidence Interval
SB	Standard Bootstrap CI
PB	Percentile Bootstrap CI
BCPB	Bias Corrected Percentile Bootstrap CI
MSE	Mean Square Error

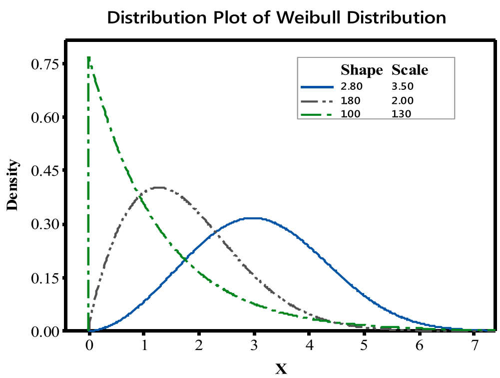


FIGURE 1. Distribution plot of Weibull distribution using different shape and scale parameters which defined low, moderate and high asymmetry.

II. METHODOLOGY

In this study, Weibull distribution with different shape and scale parameters are considered to figure out the effects of different tail behavior on PCI. The Weibull distribution with shape and scale parameters of (2.8,3.50), (1.80,2.00) and (1.00,1.30) is considered as presented in figure 1. These shape and scale parameters combinations are categorized to evaluate low, moderate and high asymmetric level of the distribution. For simulation scenario, the data sets of size $n = 25,50,75$ and 100 are generated using each asymmetric level of Weibull distribution.

A. GINI's MEAN DIFFERENCE

The Gini's mean difference for a set of n ordered observations, $\{x_1, x_2, \cdots, x_n\}$, of a random variable X is defined as

$$G_n = \frac{2}{n(n-1)} \sum_{j=1}^{n} \sum_{i=1}^{n} |x_i - x_j|$$

$$G_n = \frac{2}{n(n-1)} \sum_{i=1}^{n} [(x_i - x_1) + (x_i - x_2) + \cdots + (x_i - x_{i-1})]$$

$$G_n = \frac{2}{n(n-1)} \sum_{i=1}^{n} (2i - n - 1) x_{(i)} \tag{5}$$

If the random variable X follows normal distribution with mean μ and variance σ^2, then Downton, [22] suggests as a possible unbiased estimator of standard deviation (σ) is

$$\sigma^* = c \sum_{i=1}^{n} (2i - n - 1)\, x_{(i)} / n\,(n-1)$$

Where $c = \sqrt{\pi} = 1.77245$. Latter on, David, [23], proved that

$$\sigma^* = 0.8862 * G_n$$

is an unbiased measure of variability. GMD can be rewritten as

$$G_n = \frac{2}{n\,(n-1)} \sum_{i=1}^{n} (2i - n - 1)\, x_{(i)}$$

If we write this as

$$G_n = \frac{2}{n\,(n-1)} \sum_{i=1}^{n} ((i-1) - (n-i))\, x_{(i)}$$

$$G_n = \frac{2}{n\,(n-1)} \left[\sum_{i=1}^{n} (i-1)\, x_{(i)} - \sum_{i=1}^{n} (n-i)\, x_{(i)} \right]$$

$$G_n = \frac{2}{n\,(n-1)} [U - V]$$

Where $U = \sum_{i=1}^{n} (i-1)\, x_{(i)}$ and $V = \sum_{i=1}^{n} (n-i)\, x_{(i)}$. Using this procedure as compared to Nair [24], Lomnicki [25], the unbiased estimator of Gini's mean difference for Weibull distribution is [26],

$$E\,(G_n) = \left(2 - 2^{1-\frac{1}{\beta}}\right) \frac{\Gamma\left(1 + \frac{1}{\beta}\right)}{\lambda} = \sigma_{gw} \tag{6}$$

with p.d.f of Weibull distribution is

$$f_x = \lambda\beta\,(\lambda x)^{\beta-1}\, e^{-(\lambda x)^{\beta}} \tag{7}$$

TABLE 1. The mean and standard deviations of the Cp and Cpk using different asymmetric levels with LSL=2.0 and USL=8.0.

	Low Asymmetry		Moderate Asymmetry		High Asymmetry	
n	GMD	PC	GMD	PC	GMD	PC
			C_p			
25	1.7747 (0.2707)	1.6064 (0.2218)	1.7157 (0.2818)	1.8796 (0.3396)	2.0949 (0.4602)	1.3163 (0.4213)
50	1.8085 (0.1936)	1.5638 (0.1471)	1.7516 (0.1964)	1.8011 (0.1822)	2.1378 (0.3295)	1.2341 (0.2728)
75	1.8185 (0.1583)	1.5515 (0.1161)	1.7629 (0.1601)	1.7944 (0.1617)	2.1484 (0.2647)	1.2073 (0.2133)
100	1.8262 (0.1368)	1.5467 (0.1027)	1.7682 (0.1387)	1.7931 (0.1550)	2.1579 (0.2366)	1.1994 (0.1851)
			C_{pk}			
25	1.6953 (0.2663)	1.5362 (0.2161)	1.2480 (0.2204)	1.3739 (0.2475)	1.2318 (0.3326)	0.7625 (0.2320)
50	1.7559 (0.1979)	1.5130 (0.1453)	1.2753 (0.1562)	1.3196 (0.1656)	1.2515 (0.2305)	0.7177 (0.1522)
75	1.7696 (0.1593)	1.5141 (0.1188)	1.2818 (0.1283)	1.3132 (0.1321)	1.2563 (0.1898)	0.7058 (0.1225)
100	1.7806 (0.1388)	1.5086 (0.1013)	1.2845 (0.1106)	1.3048 (0.1177)	1.2602 (0.1657)	0.6993 (0.1055)

B. PCIs BASED ON GMD

To compute C_p and C_{pk} using GMD as a measure of variability when data follow a Weibull distribution, we have the following modification in the above non-nromal PCIS formulas.

$$C_{npg} = \frac{USL - LSL}{5.3172\sigma_{gw}} \tag{8}$$

$$C_{npkg} = \frac{\min(USL - m, m - LSL)}{2.6586\sigma_{gw}}. \tag{9}$$

C. BOOTSTRAP CONFIDENCE INTERVALS

Let $x_1, x_2, x_3, \cdots x_n$ be a random sample of size n drawn from any distribution of interest say Θ. i.e $x_1, x_2, x_3, \cdots x_n \sim \Theta$. Let $\hat{\theta}$ represents the estimator of PCI say C_p or C_{pk}. Then following steps are involved to explain the bootstrap procedure.

I. A bootstrap sample of size n (with replacement) is obtained from original sample by putting $1/n$ as mass at each point and is denoted by $x_1^*, x_2^*, x_3^* \cdots x_n^*$.

II. Let X_m^* where $1 \leq m \leq n$ be the m^{th} bootstrap sample, then m^{th} bootstrap estimator of θ is computed as

$$\hat{\theta}^* = \hat{\theta}(x_1^*, x_2^*, x_3^* \cdots x_n^*) \tag{10}$$

Where $\hat{\theta}^*$ is the m^{th} estimator of parameter $\hat{\theta}$.

III. Since there are total n^n resamples. From these resamples we calculate n^n values of $\hat{\theta}^*$. Each of these would be estimate of $\hat{\theta}$. The arrangement of the entire collection from smallest to largest, would constitute an empirical bootstrap distribution of $\hat{\theta}$.

In this study, we assumed $B = 1000$ bootstrap resamples. The construction of confidence intervals of the PCI $\hat{\theta} \epsilon \left(C_p, C_{pk}\right)$ using bootstrap techniques are described as

1) STANDARD BOOTSTRAP (SB) CONFIDENCE INTERVAL

From $B = 1000$ bootstrap estimates of $\hat{\theta}^*$, calculate the sample average and standard deviation as

$$\bar{\theta}^* = (1000)^{-1} \sum_{i-1}^{1000} \hat{\theta}^*(i) \tag{11}$$

$$S_{\hat{\theta}^*}^* = \sqrt{\left(\frac{1}{999}\right) \sum_{i=1}^{1000} \left(\hat{\theta}^*(i) - \bar{\theta}^*\right)^2} \tag{12}$$

Thus the SB $(1-\alpha)$ 100% confidence interval is

$$CI_{SB} = \bar{\theta}^* \pm Z_{1-\frac{\alpha}{2}} S_{\hat{\theta}^*}^* \tag{13}$$

Where $Z_{1-\frac{\alpha}{2}}$ is obtained by using $\left(1-\frac{\alpha}{2}\right)^{th}$ quantiles of the standard normal distribution.

2) PERCENTILE BOOTSTRAP (PB) CONFIDENCE INTERVAL

From the ordered collection of $\hat{\theta}^*(i)$, choose $100\left(\frac{\alpha}{2}\right)$ % and the $100\left(1-\frac{\alpha}{2}\right)$ % points as the end points of the confidence interval to give

$$CI_{PB} = \left(\hat{\theta}^*_{B(\frac{\alpha}{2})}, \hat{\theta}^*_{B(1-\frac{\alpha}{2})}\right) \tag{14}$$

as the $(1-\alpha)$ 100% confidence interval of $\hat{\theta}$. For a 95% confidence interval with $B = 1000$
this would be

$$CI_{PB} = \left(\hat{\theta}^*_{(25)}, \hat{\theta}^*_{(975)}\right). \tag{15}$$

3) BIAS-CORRECTED PERCENTILE BOOTSTRAP (BCPB) CONFIDENCE INTERVAL

This method has been developed to correct the potential bias. This bias is generated because the bootstrap distribution is based on a sample from the complete bootstrap distribution and may be shifted higher or lower than would be expected. The calculation of this method is based on the following steps.

i. Using the (ordered) distribution of $\hat{\theta}^*(i)$, compute the probability

$$p_0 = pr\left(\hat{\theta}^* \leq \hat{\theta}\right) \tag{16}$$

ii. Let Ø and $Ø^{-1}$ represents the cumulative and inverse cumulative distribution functions of standard normal variable z, then calculate

$$z_0 = Ø^{-1}(p_0) \tag{17}$$

iii. The percentiles of the ordered distribution of $\hat{\theta}^*$ is obtained as

$$P_L = Ø\left(2z_0 + z_{\frac{\alpha}{2}}\right) \tag{18}$$

$$P_U = Ø\left(2z_0 + z_{1-\frac{\alpha}{2}}\right) \tag{19}$$

Finally, the BCPB confidence interval is given as

$$CI_{BCPB} = \left(\hat{\theta}^*_{(P_L B)}, \hat{\theta}^*_{(P_U B)}\right). \tag{20}$$

III. RESULTS AND DISCUSSIONS

Table 1 reports the average value of GMD and PC based estimators of C_p and C_{pk} for different sample sizes under low, moderate and high asymmetric levels. The values in parenthesis are the standard deviations. Both estimators performed differently under different tail behavior of the Weibull distribution. In all asymmetric levels, GMD based estimators of C_p perform better than its competitor. However, under low and moderate asymmetry it is very close to target value and produce lower bias and Mean Square Error (MSE). Although, PC based estimator of C_p is good up to some extent but not recommended for new processes. However, it may produce better results for existing processes where Cp=1.33 or higher. In case of high asymmetry, however, the efficiency of both estimators differ significantly. The quantile approach is not good and exhibits lower values indicating that process does not meet the specification limits.

The performance of GMD based estimator of Cp is more robust and gives higher values under high asymmetry. It has shown lower bias and MSE when Cp values equal to 2.00 or higher. In case of Cpk index, the performance of both estimators was quite different, however both underestimate the true Cpk under moderate and high asymmetry, although the performance of GMD is clearly better than PC. Under low asymmetry GMD based estimator has given reliable results.

The MSE under low, moderate and high asymmetry using different sample sizes and standard values of C_p and C_{pk} are presented in figure 2 - 7 respectively using radar chart. From these charts, it is concluded that MSE in case of the GMD based estimator is less than PC based estimator under all asymmetric levels.

Under the same simulation setup for point estimation, the confidence interval and their coverage probabilities for both estimators are listed in table 2–5. From these tables, it is concluded that coverage probabilities are increasing and average widths are decreasing when the sample size increasing. Moreover, it is concluded that BCPB

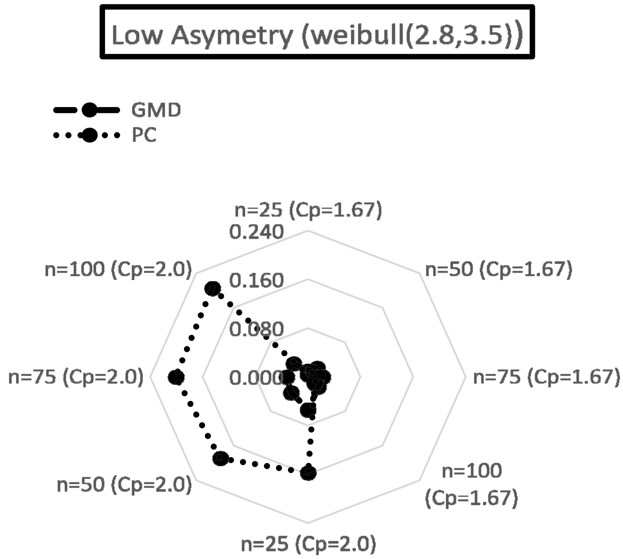


FIGURE 2. MSE under low asymmetry for Cnpg index.

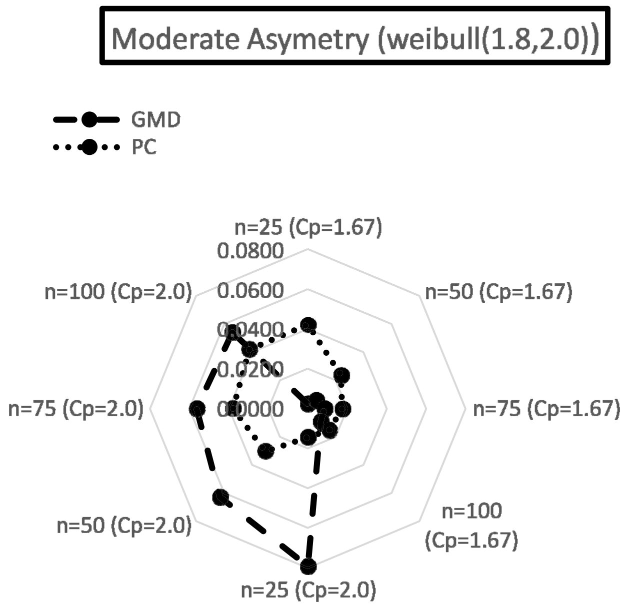


FIGURE 3. MSE under moderate asymmetry for Cnpg index.

method has the highest coverage probabilities under all asymmetric levels using GMD method. On the other hand, SB method performed batter using quantile approach for both indices. The coverage probabilities in both cases reaches the nominal confidence coefficient 0.95 using large sample sizes.

The results show that, under all asymmetric level, with $n \geq 50$ all three bootstrap methods provides enough coverage proportions and reaches the nominal confidence coefficient 0.95. In case of GMD based estimator, SB and PB have lower coverage proportions while in case of quantiles PB and BCBP provides poor coverage proportions. The performance of three bootstrap confidence intervals

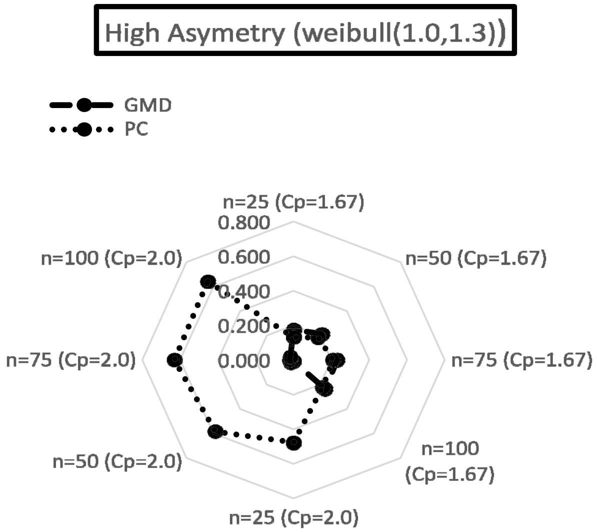


FIGURE 4. MSE under high asymmetry for Cnpg index.

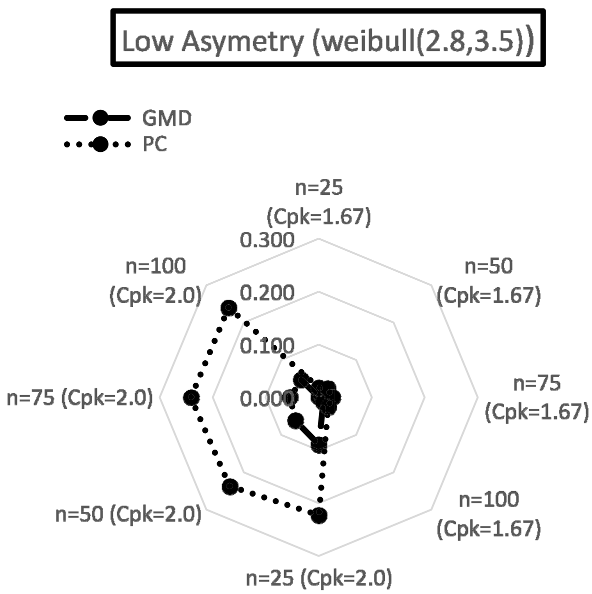


FIGURE 5. MSE under low asymmetry for Cnpkg index.

based on lower average widths using GMD are ranked as $SB > BCPB > PB$. While in case of PC, it has following rank $SB > PB > BCPB$ for index C_p. On the other hand, in case of index, C_{pk}, we observed the following order $BCPB > SB > PB$ and $SB > PB > BCBP$ using GMD and quantile estimators respectively. Therefore, based on better coverage probabilities and lower average widths, BCBP confidence limits are reliable for index C_p using both approaches and for C_{pk} using quantile approach only. The PB method provides lower confidence limits using GMD based estimator of C_{pk}.

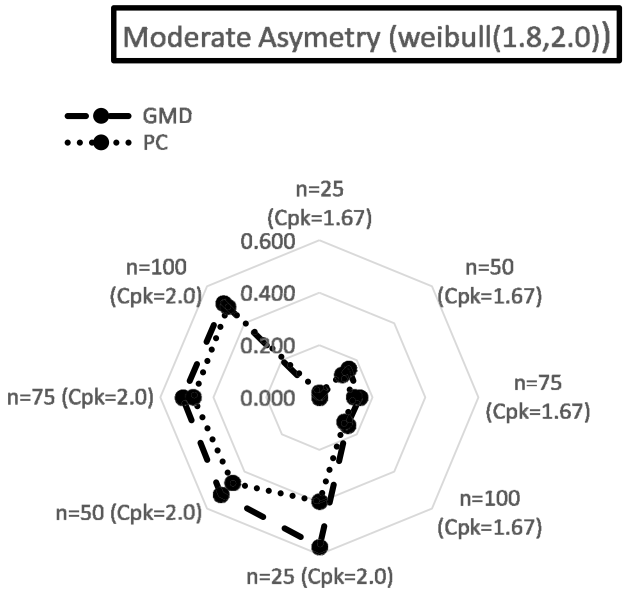


FIGURE 6. MSE under moderate asymmetry for Cnpkg index.

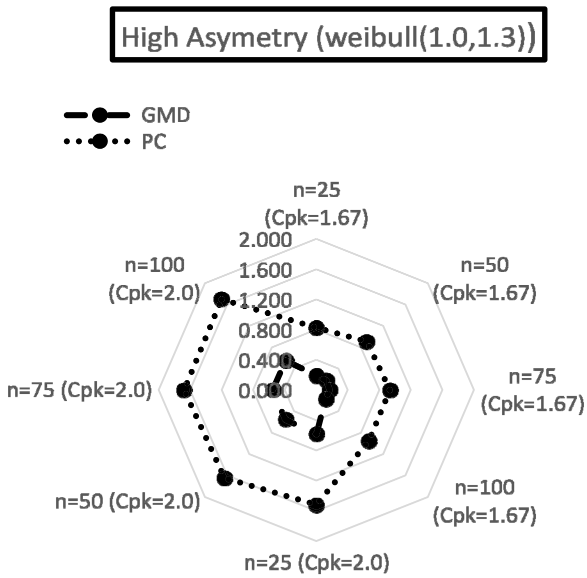


FIGURE 7. MSE under high asymmetry for Cnpkg index.

A. EXAMPLE

The manufacturing data of floor tiles is taken from [27]. The company is concerned about the flexibility of the tiles, and the data set contains data collected on 10 tiles produced on each of 10 consecutive working days. Suppose that a tile manufacturer needs to keep the degree of warping in a ceramic bath tile between 2 and 8 millimeters. The basic normality test confirms that normal distribution does not model the data well. So we cannot trust on the results of standard PCIs. The goodness of fit test indicate that Weibull distribution is a good for this data. So we performed a non-normal process capability analysis. The shape and scale parameters for this data are 1.69368 and 3.27812 respectively. The low value of

TABLE 2. Three Bootstrap CIs in parentheses and their coverage probabilities for index C_{npg} using GMD method.

n	Cp	SB	PB	BCPB
Low Asymmetry				
25	1.84	(1.2145-2.2278) 0.8600	(1.2232-2.2165) 0.8650	(1.3521-2.3482) **0.9180**
50	1.84	(1.1990-1.8596) 0.9130	(1.2190-1.8652) 0.9110	(1.2946-1.9310) **0.9370**
75	1.84	(1.5397-2.1430) 0.9340	1.5475-2.1639) 0.9360	(1.5703-2.2122) **0.9510**
100	1.84	(1.5622-2.0872) 0.9350	(1.5601-2.0977) 0.9380	(1.5878-2.1286) **0.9510**
Moderate Asymmetry				
25	1.78	(1.0887-2.0792) 0.8450	(1.1107-2.0901) 0.8540	(1.2521-2.2652) **0.9160**
50	1.78	(1.1454-1.8001) 0.9060	(1.1549-1.8029) 0.9080	(1.2352-1.8665) **0.9450**
75	1.78	(1.4789-2.0994) 0.9320	(1.4942-2.1214) 0.9290	(1.5175-2.1439) **0.9520**
100	1.78	(1.4993-2.0298) 0.9390	(1.5125-2.0350) 0.9370	(1.5240-2.0565) **0.9470**
High Asymmetry				
25	2.18	(0.8289-2.4702) 0.8480	(0.8769-2.4913) 0.8630	(1.0919-2.8338) **0.9220**
50	2.18	(1.2739-2.2708) 0.9140	(1.2897-2.2871) 0.9160	1.3886-2.3884) **0.9390**
75	2.18	(1.6802-2.7383) 0.9440	(1.6936-2.7518) 0.9440	(1.7503-2.8007) **0.9530**
100	2.18	(1.7125-2.5834) 0.9250	(1.7384-2.6185) 0.9320	1.7537-2.6413) **0.9360**

TABLE 3. Three Bootstrap CIs in parentheses and their coverage probabilities for index C_{npg} using PC method.

n	Cp	SB	PB	BCPB
Low Asymmetry				
25	1.53	(1.2386-2.1434) 0.9220	(1.3106-2.1906) 0.8600	(1.2285-2.0390) **0.9200**
50	1.53	(1.4735-2.1182) 0.9420	(1.5098-2.1448) 0.9120	(1.4720-2.0419) **0.9440**
75	1.53	(1.3152-1.7505) 0.9520	(1.3216-1.7616) 0.9340	(1.3025-1.7373) **0.9500**
100	1.53	(1.3571-1.7378) 0.9530	(1.3716-1.7429) 0.9390	(1.3537-1.7342) **0.9510**
Moderate Asymmetry				
25	1.77	(1.3108-2.1891) 0.9180	(1.4402-2.9030) 0.8600	(1.3139-2.6228) **0.9170**
50	1.77	(1.6774-2.6667) 0.9390	(1.7419-2.7156) 0.9120	(1.6881-2.5893) **0.9420**
75	1.77	(1.4274-2.1117) 0.9530	1.4464-2.1313) 0.9330	(1.4251-2.1045) **0.9510**
100	1.77	(1.4970-2.0944) 0.9530	(1.5173-2.1163) 0.9370	(1.5103-2.0917) **0.9480**
High Asymmetry				
25	1.16	(0.5351-2.6018) 0.9330	(0.8147-2.7716) 0.8640	(0.6934-2.3343) **0.9140**
50	1.16	(1.0182-2.3094) 0.9500	(1.1289-2.3709) 0.9140	(1.0772-2.2394) **0.9480**
75	1.16	(0.7746-1.5701) 0.9640	(0.8155-1.6111) 0.9310	(0.7999-1.5832) **0.9480**
100	1.16	(0.8528-1.5515) 0.9550	(0.8883-1.5959) 0.9410	(0.8821-1.5651) **0.9480**

TABLE 4. Three Bootstrap CIs in parentheses and their coverage probabilities for index C_{npkg} using GMD method.

n	Cp	SB	PB	BCPB
Low Asymmetry				
25	1.82	(1.1423-2.1074) 0.8090	(1.1658-2.1252) **0.8270**	(1.3373-2.2842) 0.9020
50	1.82	(1.1501-1.8176) 0.8860	(1.1765-1.8196) **0.8900**	(1.2747-1.9013) 0.9260
75	1.82	(1.4802-2.0971) 0.9170	(1.4826-2.1200) **0.9200**	(1.5585-2.2013) 0.9500
100	1.82	(1.5163-2.0513) 0.9270	(1.5104-2.0676) **0.9310**	(1.5694-2.1095) 0.9360
Moderate Asymmetry				
25	1.29	(0.7043-1.4504) 0.8450	(0.7281-1.4624) **0.8570**	(0.8264-1.6168) 0.9240
50	1.29	(0.8256-1.3293) 0.9150	(0.8351-1.3416) **0.9170**	(0.8776-1.3787) 0.9390
75	1.29	(1.0547-1.5608) 0.9390	(1.0600-1.5685) **0.9410**	(1.0918-1.5897) 0.9520
100	1.29	(1.0699-1.4895) 0.9260	(1.0814-1.5048) **0.9350**	(1.0900-1.5174) 0.9470
High Asymmetry				
25	1.27	(0.3730-1.4211) 0.8780	(0.4343-1.4869) **0.8910**	(0.5132-1.6654) 0.9310
50	1.27	(0.6981-1.3862) 0.9240	(0.7074-1.4004) **0.9280**	(0.7620-1.4855) 0.9380
75	1.27	(0.9155-1.6780) 0.9450	(0.9409-1.6888) **0.9490**	(0.9559-1.7264) 0.9540
100	1.27	(0.9402-1.5543) 0.9260	(0.9572-1.5704) **0.9270**	(0.9573-1.5758) 0.9370

TABLE 5. Three Bootstrap CIs in parentheses and their coverage probabilities for index C_{npkg} using PC method.

n	Cp	SB	PB	BCPB
Low Asymmetry				
25	1.51	(1.1579-2.0361) 0.9470	(1.2131-2.0877) 0.9180	(1.1631-1.9990) **0.9300**
50	1.51	(1.4188-2.0656) 0.9610	(1.4497-2.0731) 0.9510	(1.4345-2.0463) **0.9460**
75	1.51	(1.2661-1.7114) 0.9610	(1.2803-1.7268) 0.9530	(1.2891-1.7374) **0.9500**
100	1.51	(1.3191-1.7062) 0.9560	(1.3222-1.7160) 0.9540	(1.3512-1.7462) **0.9470**
Moderate Asymmetry				
25	1.28	(0.8938-1.9075) 0.9090	(0.9702-1.9789) 0.8710	(0.8847-1.7942) **0.9170**
50	1.28	(1.2258-1.9505) 0.9370	(1.2668-1.9584) 0.9170	(1.2111-1.8502) **0.9370**
75	1.28	(1.0431-1.5413) 0.9500	(1.0469-1.5543) 0.9370	(1.0279-1.5350) **0.9540**
100	1.28	(1.0812-1.5224) 0.9500	(1.0929-1.5410) 0.9390	(1.0754-1.5146) **0.9500**
High Asymmetry				
25	0.68	(0.3061-1.3755) 0.9230	(0.4427-1.4638) 0.8590	(0.3781-1.2542) **0.9180**
50	0.68	(0.6052-1.3426) 0.9430	(0.6581-1.3796) 0.9080	(0.6180-1.2556) **0.9420**
75	0.68	(0.4583-0.9120) 0.9570	(0.4760-0.9333) 0.9330	(0.4516-0.9039) **0.9500**
100	0.68	(0.4951-0.8974) 0.9570	(0.5107-0.9168) 0.9320	(0.4941-0.8947) **0.9530**

the shape parameter indicates that the data is right skewed. The summary statistics of the data and calculated values of both indices C_{npg} and C_{npkg} were presented in Table 6.

The process capability analysis of the data indicate that process is not being capable when standard deviation and quantile approach is used as a measure of variability. However, when GMD is used the process is being capable (Cp $>$ 1.33). Usually all companies consider Cp instead of Cpk for evaluating their manufacturing processes. Table 6 indicates that the production process is not centered in relation to the specification limits because median is less than average of upper and lower specification limits. For this reason, we must consider Cpk along with Cp. Using Cpk, the situation is somewhat different to that observed in using Cp. The results indicate that for positively skewed data, all methods underestimate the actual process yield. However, using GMD one would be able to reduce the nonconforming parts because the value of modified Cpk is much greater than the qantile based Cpk. Three bootstrap confidence intervals of the both PCs and their coverage probabilities are reported in table (7). The true values of both indices lie in the bootstrap confidence intervals. Additionally, the results are similar to

TABLE 6. Summary statistics of the data.

Statistics	Value
Min.	0.2819
Max.	8.0910
Mean	2.9231
S.d	1.7859
Q(0.00135)	0.0663
Q(0.05)	2.6402
Q(0.99865)	9.9955
GMD	0.5789
Cp-Standard	0.5900
Cp-PC	0.6000
Cp-GMD	1.7300
Cpk-Standard	0.1700
Cpk-PC	0.1300
Cpk-GMD	0.3700

TABLE 7. Bootstrap CIs with their coverage probabilities using both methods for Cp and Cpk.

Method	SB	PB	BCPB
		Cp	
GMD	(1.4493-1.9684) 0.9400	(1.4614-1.9796) 0.9340	(1.4736-1.9886) 0.9470
PC	(0.5069-0.7224) 0.9530	(0.5141-0.7308) 0.9380	(0.5120-0.7222) 0.9480
		Cpk	
GMD	(0.1483-0.5716) 0.9450	(0.1647-0.5849) 0.9460	(0.1647-0.5849) 0.9430
PC	(0.0564-0.2012) 0.9580	(0.0616-0.2001) 0.9580	(0.0500-0.1935) 0.9580

the simulation results.

IV. CONCLUSION AND RECOMMENDATIONS

Considering different skewed Weibull processes, this study proposes GMD based process capability indices. The GMD is used to measure the process variability as compared to standard deviation to evaluate the performance of C_p and C_{pk}. For point estimation, GMD based PCIs, has lower MSE under all asymmetric levels as compared to PC based PCIs. The GMD based PCIs recommend for new processes where the target value is not less than 1.67. On the other hand, The PC- based PCIs performed well under low asymmetry for just capable processes. It is also observed that at high asymmetry, GMD based PCIs are more efficient than PC based PCIs. Even under low and moderate asymmetry, GMD is clearly non-inferior to its competitor.

The major advantage of applying GMD philosophy is that it helpful for the reduction of process variability under high asymmetry and process can meet the customer's requirement. Further, we focused attention on deriving non-parameters confidence intervals for both GMD and PC based PCIs under low, moderate and high asymmetry. In case of C_p, the BCBP methods provide reliable confidence limits and better coverage probability using GMD and quantile methods for underlying asymmetric levels whatever the sample size. The PB method provides higher coverage probability with smaller confidence interval width in case of C_{pk} using Ginni's mean difference method.

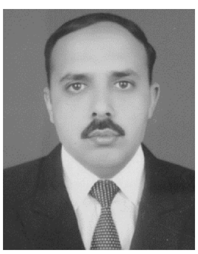
ACKNOWLEDGMENT

The authors are deeply thankful to editor and two reviewers for their valuable suggestions to improve the quality of manuscript. The authors, therefore, acknowledge with thanks DSR technical and financial support.

REFERENCES

[1] S. Safdar and E. Ahmed, "Process capability indices for shape parameter of weibull distribution," *Open J. Statist.*, vol. 4, no. 3, pp. 207–219, 2014.
[2] M. R. Piña-Monarrez, J. F. Ortiz-Yañez, and M. I. Rodríguez-Borbón, "Non-normal capability indices for the weibull and lognormal distributions," *Quality Rel. Eng. Int.*, vol. 32, no. 4, pp. 1321–1329, Jul. 2015.
[3] J. M. Juran, *Jurans Quality Control Handbook*, 3rd ed. New York, NY, USA: McGraw-Hill, 1974.
[4] V. E. Kane, "Process capability indices," *J. Quality Technol.*, vol. 18, no. 1, pp. 41–52, 1986.
[5] O. Senvar and B. Sennaroglu, "Comparing performances of clements, box-cox, Johnson methods with weibull distributions for assessing process capability," *J. Ind. Eng. Manage.*, vol. 9, no. 3, p. 634, 2016.
[6] B. Sennaroglu and O. Senvar, "Performance comparison of box-cox transformation and weighted variance methods with weibull distribution," *J. Aeronautics Space Technol.*, vol. 8, pp. 49–55, Sep. 2015.
[7] S.-C. Kao, "Deciding optimal specification limits and process adjustment under quality loss function and process capability indices," *Int. J. Ind. Eng.*, vol. 17, no. 3, pp. 212–222, 2010.
[8] O. Senvar and C. Kahraman, "Type-2 fuzzy process capability indices for non-normal processes," *J. Intell. Fuzzy Syst.*, vol. 27, no. 2, pp. 769–781, 2014.
[9] O. Senvar and C. Kahraman, "Fuzzy process capability indices using Clements' method for non-normal processes," *J. Multiple-Valued Logic Soft Comput.*, vol. 22, nos. 1–2, pp. 95–121, 2014.
[10] L. C. Tang and S. E. Than, "Computing process capability indices for non-normal data: A review and comparative study," *Quality Rel. Eng. Int.*, vol. 15, no. 5, pp. 339–353, Sep./Oct. 1999.
[11] J. A. Clements, "Process capability calculations for non-normal distributions," *Quality Prog.*, vol. 22, pp. 95–100, Sep. 1989.
[12] W. Pearn and K. Chen, "Capability indices for non-normal distributions with an application in electrolytic capacitor manufacturing," *Microelectron. Rel.*, vol. 37, no. 12, pp. 1853–1858, 1997.
[13] C. W. Wu, C. S. Chang, W. L. Pearn, and H. C. Chen, "Accuracy analysis of the percentile method for estimating non normal manufacturing quality," *Commun. Statist.-Simul. Comput.*, vol. 36, no. 3, pp. 657–696, 2007.
[14] M. Kovářík and L. Sarga, "Process capability indices for non-normal data," *WSEAS Trans. Bus. Econ.*, vol. 11, no. 1, pp. 419–429, 2014.
[15] D. C. Nanthakumar and M. S. V. Vijayalakshmi, "Construction of interquartile range (IQR) control chart using process capability for mean," *Int. J. Modern Sci. Eng. Technol.*, vol. 5, no. 1, pp. 114–118, 2015.
[16] S. Yitzhaki, "Gini's mean difference: A superior measure of variability for non-normal distributions," *Metron*, vol. 61, no. 2, pp. 285–316, 2003.
[17] M. Zenga, M. Polisicchio, and F. Greselin, "The variance of Gini's mean difference and its estimators," *Statistica*, vol. 64, no. 3, pp. 455–475, 2004.
[18] A. Čiginas and D. Pumputis, "Gini's mean difference and variance as measures of finite populations scales," *Lithuanian Math. J.*, vol. 55, no. 3, pp. 312–330, 2015.
[19] C. Gini, *Variabilita e Mutabilita, Contributo Allo Studio Delle Distribuzioni e Delle Relazioni Statistiche*, vol. 3. Cagliari: Universita di Cagliari, 1912, pp. 153–159.
[20] M. Riaz and A. Saghirr, "Monitoring process variability using Gini's mean difference," *Quality Technol. Quant. Manage.*, vol. 4, no. 4, pp. 439–454, 2007.

[21] C. Gerstenberger and D. Vogel, "On the efficiency of Gini's mean difference," *Statist. Methods Appl.*, vol. 24, no. 4, pp. 569–596, 2015.
[22] F. Downton, "Linear estimates with polynomial coefficient," *Biometrika*, vol. 53, no. 1/2, pp. 129–141, 1966.
[23] H. David, "Miscellanea: Gini's mean difference rediscovered," *Biometrika*, vol. 55, no. 3, pp. 573–575, 1968.
[24] U. Nair, "The standard error of Gini's mean difference," *Biometrika*, vol. 28, no. 3/4, pp. 428–436, 1936.
[25] Z. A. Lomnicki, "The standard error of Gini's mean difference," *The Ann. Math. Statist.*, vol. 23, no. 4, pp. 635–637, 1952.
[26] P. Huang and T. Hwang, "The inference of Gini's mean difference," *Int. J. Pure Appl. Math.*, vol. 25, no. 1, pp. 39–48, 2005.
[27] *MINITAB Release 17: Statistical Software for Windows*, Minitab Inc., State College, PA, USA, 2014.



MUHAMMAD KASHIF received the B.S. and M.S. degrees in statistics from the University of Agriculture, Faisalabad, Pakistan, in 2002 and 2009, respectively. He is currently pursuing the Ph.D. degree in statistics from King Abdulaziz University, Jeddah, Saudi Arabia, under its Fellowship Programme. He has been a Lecturer with the Department of Statistics, University of Agriculture, since 2006. He has authored or co-authored over 15 technical papers in applied Statistics. His research interests include the nonnormal distribution, process capability indices, time series analysis, and the design of experiments.

MUHAMMAD ASLAM received the M.Sc. degree in statistics, with the Chief Minister of the Punjab Merit Scholarship, and the M.Phil. degree in statistics, with the Governor of the Punjab Merit Scholarship, from GC University, Lahore, in 2004 and 2006, respectively, and the Ph.D. degree in statistics from the National College of Business Administration and Economics, Lahore, in 2010, under the supervision of Dr. M. Ahmad. He was a Lecturer of Statistics with Edge College System International from 2003 to 2006. He was a Research Assistant with the Department of Statistics, GC University, from 2006 to 2008. Then, he joined the Forman Christian College University as a Lecturer in 2009, where he was an Assistant Professor from 2010 to 2012. He was an Associate Professor with the Department of Statistics from 2012 to 2014. He is currently an Associate Professor of Statistics with the Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia. He has authored over 185 research papers in national and international journals, including the IEEE Access, *Journal of Applied Statistics*, *European Journal of Operation Research*, *Journal of the Operational Research Society*, *Applied Mathematical Modeling*, *International Journal of Advanced Manufacturer Technology*, *Communications in Statistics*, *Journal of Testing and Evaluation*, and *Pakistan Journal of Statistics*. He has authored one book published in Germany. He has been also an HEC approved Ph.D. Supervisor since 2011. He supervised four Ph.D. theses, over 25 M.Phil. theses, and three M.Sc. theses. He is currently supervising two Ph.D. theses and over five M.Phil. theses in statistics. He reviewed over 70 research papers for various well reputed international journals. He is a member of Editorial Board of *Electronic Journal of Applied Statistical Analysis* and *Pakistan Journal of Commence and Social Sciences*. His areas of interest include reliability, decision trees, industrial statistics, acceptance sampling, rank set sampling, and applied statistics. He is a member of the Islamic Countries Society of Statistical Sciences. He received a Meritorious Services Award in research from the National College of Business Administration and Economics in 2011. He received the Research Productivity Award for the year 2012 by Pakistan Council for Science and Technology. His name listed at second position among Statistician in the Directory of Productivity Scientists of Pakistan in 2013. His name listed at first position among Statistician in the Directory of Productivity Scientists of Pakistan in 2014. He got 371st position in the list of top 2210 profiles of Scientist of Saudi Institutions in 2016.

ALI HUSSEIN AL-MARSHADI received the M.S. degree in statistics from New Mexico State University, La Cruces, NM, USA, in 1997 and the Ph.D. degree in statistics from Oklahoma State University, Stillwater, OK, USA, in 2004. In 1990, he joined the Faculty of King Abdulaziz University, Jeddah, Saudi Arabia, where he is currently a Professor with the Department of Statistics. He has authored or co-authored over 20 technical papers in different areas of statistics. His current research interests include experimental design, linear models, sampling, data mining, process capability indicies and statistical modeling and simulation.

CHI-HYUCK JUN was born in Seoul, South Korea, in 1954. He received the B.S. degree in mineral and petroleum engineering from Seoul National University, in 1977, the M.S. degree in industrial engineering from KAIST, in 1979, and the Ph.D. degree in operations research from the University of California, Berkeley, in 1986. Since 1987, he has been with the Department of Industrial and Management Engineering, Pohang University of Science and Technology, where he is currently a Professor and the Head of the Department of Industrial and Management Engineering. His current research interests include reliability and quality analysis and data mining techniques. He is a member of the INFORMS and the ASQ.