



MODIFICATION OF CLEMENTS' METHOD FOR ASSESSING THE CAPABILITY OF A NON-NORMAL PROCESS WITH AN APPLICATION

Burcu AYTAÇOĞLU^{1,*}, Cemil Giray GENÇ²

¹ Department of Statistics, Faculty of Science, Ege University, İzmir, Turkey

² Department of Statistics, Faculty of Science, Ege University, İzmir, Turkey

ABSTRACT

In many industries process capability studies are conducted in order to determine the capability of the process to produce acceptable products and it is one of the important activities of statistical process control. In order to express the capability of a process, process capability indices are frequently used. However, they are usually computed under the assumption that the process data follow a normal distribution. Normality assumption may be violated and the use of these traditional capability indices may cause misleading interpretation about the capability of the process. One of the most widely discussed methods to handle non-normality is Clements' method which was proposed in 1989. Clements' method uses the Pearson family of curves for calculating capability indices for any shape of distribution. It requires the estimation of the mean, standard deviation, skewness and kurtosis and makes use of the classical estimators of skewness and kurtosis. In this study, we discussed the use of more robust estimators of skewness and kurtosis in the calculation of process capability index by Clements' method. For this purpose, capability indices with the use of these robust estimators are computed by simulation and the mean square errors of them are reported. The comparison is done through simulating Weibull and lognormal data with several different parameter values. Finally, a real life application to oil pump manufacturing in automotive industry is presented.

Keywords: Process Capability Index, Non-normality, Statistical Process Control, Skewed Population, Robust Estimator

1. INTRODUCTION

Process capability analysis is the activity of analyzing process variability relative to product specifications and it helps manufacturing in decreasing variability and improving quality. Process capability indices (PCIs) are frequently used to quantify the capability of the process in producing items within the specifications. The most widely used indices in industry; C_p and C_{pk} were defined as

$$C_p = \frac{USL - LSL}{6\sigma} \quad (1)$$

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

where LSL and USL denote the lower and upper specification limits, respectively. Since the process mean μ and variance σ^2 are unknown they are estimated from the sample.

The traditional PCIs are usually computed and interpreted under the assumption that the process data follow a normal distribution. However, normality assumption may be violated and the use of these indices may cause misleading interpretation about the capability of the process.

Several studies are available in literature that account for non-normality of the process data. There are usually two ways to handle a non-normal process data. The one is to transform non normal data to normal data and to use standard PCIs given in (1) and (2). Numerous normalizing transformations were proposed in the literature to use for this purpose. (see, for example, Box and Cox [1], Johnson and Kotz [2], Ryan [3], Ryan and Schwertman [4], Niaki and Abbasi [5]). The other is to use PCIs defined directly for non-normal data. Clements [6] used the Pearson family of curves for calculating capability indices for any

*Corresponding Author: burcu.aytacoglu@ege.edu.tr

Received: 17.01.2019 Published: 30.12.2019

shape of distribution. Liu and Chen [7] have modified Clements’ method using the Burr XII distribution. Wu et al. [8] proposed weighted variance PCIs which are based on the semivariance approximation of Choobineh and Branting [9]. Chang et.al.[10] computed PCIs which are based on weighted standard deviation method. In addition, there are also several comparative studies for the proposed methods of PCI calculation.(see, for example, Tang and Than [11], Kovarik and Sarga [12], Sennaroglu and Senvar [13], Hosseinifard et al. [14]). Authors may also benefit from the review papers by Kotz and Johnson [15], Anis [16], Wu et al. [17].

In this study, we considered Clements’ method [6] which is the most widely discussed methods to handle non-normality in capability index calculations. It requires the estimation of the mean, standard deviation, skewness and kurtosis and makes use of the classical measures of skewness and kurtosis. We modified the Clements’ method by using the alternative measures of skewness and kurtosis some of which are more robust than the classical measures. These are Bowley coefficient of skewness (sk_2), Groeneveld and Meeden coefficient of skewness (sk_3), Pearson coefficient of skewness (sk_4), Moors coefficient of kurtosis (kr_2), the centered Hogg coefficient of kurtosis (kr_3), and the centered coefficient of kurtosis suggested by Crow and Siddiqui [18] (kr_4). They are briefly discussed in Section 3. The robustness of these measures was studied by Kim and White [19]. They investigated the robustness of the measures by adding a single outlier to the generated data and also by considering mixture distribution. We use the measures which are available in Kim and White [19] in the context of Clements’ method. For comparative purposes, Weibull and lognormal data with several different parameter values were generated and mean square errors (MSEs) of the estimators of capability indices were simulated. These distributions were considered since they are frequently encountered in industrial processes when the process is not symmetric and they are the most discussed asymmetric distributions in process capability studies.

This article is organized as follows. In Section 2, the details of the Clements’ method is explained. In Section 3, the measures of skewness and kurtosis which are used in this study are discussed. In Section 4, the simulation procedure is explained and the MSEs are reported. In Section 5, a real life application to oil pump manufacturing in automotive industry is presented and some concluding remarks are given in Section 6.

2. CLEMENTS’ METHOD

Clements [6] used the Pearson family of curves for calculating capability indices for any shape of distribution. In the Clements’ method 6σ in (1) was replaced by $U_p - L_p$, giving

$$C_p = \frac{USL-LSL}{U_p-L_p} \tag{3}$$

where U_p and L_p denote 99.865 and 0.135 percentiles, respectively. The process mean μ in (2) was estimated by the median M , and 3σ was replaced by $U_p - M$ for finding C_{pu} and by $M - L_p$ for finding C_{pl} , giving

$$C_{pu} = \frac{USL-M}{U_p-M} \tag{4}$$

$$C_{pl} = \frac{M-LSL}{M-L_p} \tag{5}$$

C_{pk} is defined as the minimum of C_{pu} and C_{pl} as follows:

$$C_{pk} = \min \left\{ \frac{USL-M}{U_p-M}, \frac{M-LSL}{M-L_p} \right\} \tag{6}$$

2.1. Procedure for Calculating PCIs by Clements’ Method

In the calculation of PCIs by Clements’ method, estimates of the mean, standard deviation, skewness and kurtosis of the process data are required, and they should be calculated from a stable, in control process. USL and LSL are usually given externally by the design engineers according to the design requirements of the product.

The following steps give the procedure for calculating PCIs by Clements’ Method.

1. LSL and USL are known values since they are determined by the design engineers.
2. The process mean and standard deviation are estimated by \bar{X} and S, respectively.
3. Skewness (sk) and kurtosis (kr) are estimated by the traditional method which is based on the third and the fourth central moments. (Detailed formulas are given in section 3)
4. Standardized values of L_p , U_p and M (denoted by L'_p , U'_p and M') are obtained from the tables provided by [6].
5. L_p , U_p and M are estimated as follows

$$L_p = \bar{x} - sL'_p, \quad U_p = \bar{x} + sU'_p, \quad M = \bar{x} + sM'$$

6. PCIs are calculated from equations (3)-(6).

As mentioned before, the given procedure for calculating PCIs employs traditional estimators of skewness and kurtosis. In this study, the use of more robust estimators of skewness and kurtosis in the calculation of PCIs are considered and an extensive simulation work is conducted for comparison. Skewness and kurtosis measures used in this study are explained in detail in section 3.

3. SKEWNESS AND KURTOSIS MEASURES

Suppose that observations are obtained from a stable process and they are denoted by $\{y_i\}$, $i=1,2,\dots,n$. Moreover, y_i 's are identically, independently distributed and have the cumulative distribution function of F, and F^{-1} stands for the inverse of the cumulative distribution function. The traditional coefficients of skewness and kurtosis for y_i are given by

$$sk_1 = E \left(\frac{y_i - \mu}{\sigma} \right)^3 \tag{7}$$

$$kr_1 = E \left(\frac{y_i - \mu}{\sigma} \right)^4 - 3 \tag{8}$$

where $\mu = E(y_i)$ and $\sigma = E(y_i - \mu)^2$. The following estimators are usually used in order to estimate sk_1 and kr_1 .

$$\widehat{sk}_1 = n^{-1} \sum_{i=1}^n \left(\frac{y_i - \hat{\mu}}{\hat{\sigma}} \right)^3 \tag{9}$$

$$\widehat{kr}_1 = n^{-1} \sum_{i=1}^n \left(\frac{y_i - \hat{\mu}}{\hat{\sigma}} \right)^4 - 3 \tag{10}$$

where $\hat{\mu} = n^{-1} \sum_{i=1}^n y_i$ and $\hat{\sigma} = n^{-1} \sum_{i=1}^n (y_i - \hat{\mu})^2$.

\widehat{sk}_1 and \widehat{kr}_1 become arbitrarily large when there are one or more large outliers in the data. Therefore, some other measures of skewness and kurtosis that are less sensitive to the outliers were proposed in the literature. For example, Bowley [20] proposed a coefficient of skewness based on quantiles

$$sk_2 = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} \tag{11}$$

where Q_k is the k^{th} quartile of y_i ; that is, $Q_1 = F^{-1}(0.25)$, $Q_2 = F^{-1}(0.5)$, and $Q_3 = F^{-1}(0.75)$. The estimator of sk_2 is given by

$$\widehat{sk}_2 = \frac{\widehat{Q}_3 + \widehat{Q}_1 - 2\widehat{Q}_2}{\widehat{Q}_3 - \widehat{Q}_1} \tag{12}$$

where \widehat{Q}_k is the k^{th} sample quartile.

The Bowley coefficient of skewness has been generalized by Hinkley [21],

$$sk_3(\alpha) = \frac{F^{-1}(1-\alpha) + F^{-1}(\alpha) - 2Q_2}{F^{-1}(1-\alpha) - F^{-1}(\alpha)} \tag{13}$$

for any α between 0 and 0.5.

In order to remove the dependence of this measure on the value of α , Groeneveld and Meeden [22] integrated out α as:

$$sk_3 = \frac{\int_0^{0.5} \{F^{-1}(1-\alpha) + F^{-1}(\alpha) - 2Q_2\} d\alpha}{\int_0^{0.5} \{F^{-1}(1-\alpha) - F^{-1}(\alpha)\} d\alpha} = \frac{\mu - Q_2}{E|y_t - Q_2|} \tag{14}$$

and this coefficient of skewness is estimated by

$$\widehat{sk}_3 = \frac{\widehat{\mu} - \widehat{Q}_2}{n^{-1} \sum_{i=1}^n |y_i - \widehat{Q}_2|} \tag{15}$$

The final measure of skewness used in this study is the Pearson coefficient [23], and it was obtained by replacing the denominator in sk_3 with the standard deviation.

$$sk_4 = \frac{\mu - Q_2}{\sigma} \tag{16}$$

The following is the estimator of sk_4 .

$$\widehat{sk}_4 = \frac{\widehat{\mu} - \widehat{Q}_2}{\widehat{\sigma}} \tag{17}$$

Moors [24] proposed an alternative measure of kurtosis which was given by

$$kr'_2 = \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2} \tag{18}$$

where E_k is the k^{th} octile, i.e. $E_k = F^{-1}(k/8)$ for $k=1, 2, \dots, 7$.

Moors coefficient of kurtosis is centered at the value for the standard normal distribution which is equal to 1.23 [19]. Therefore, the centered coefficient was given by

$$kr_2 = \frac{(E_7 - E_5) + (E_3 - E_1)}{E_6 - E_2} - 1.23 \tag{19}$$

and the estimator of kr_2 is

$$\widehat{kr}_2 = \frac{(\widehat{E}_7 - \widehat{E}_5) + (\widehat{E}_3 - \widehat{E}_1)}{\widehat{E}_6 - \widehat{E}_2} - 1.23 \tag{20}$$

where \widehat{E}_k is the k^{th} sample octile.

Hogg [25, 26] suggested another measure of kurtosis which is given by $\frac{U_\alpha - L_\alpha}{U_\beta - L_\beta}$ where $U_\alpha(L_\alpha)$ is the average of the upper(lower) α quantiles defined as:

$$U_\alpha = \int_{1-\alpha}^1 F^{-1}(y) dy, \quad L_\alpha = \int_0^\alpha F^{-1}(y) dy, \quad \text{for } \alpha \in (0,1)$$

According to Hogg’s simulation experiments, $\alpha = 0.05$ and $\beta = 0.5$ gave the most satisfactory results. Adopting these values for α and β , Hogg coefficient was calculated as 2.59 for standard normal distribution [19]. Hence, the centered Hogg coefficient was given by

$$kr_3 = \frac{U_{0.05} - L_{0.05}}{U_{0.5} - L_{0.5}} - 2.59 \tag{21}$$

The estimator of kr_3 is

$$\widehat{kr}_3 = \frac{\widehat{U}_{0.05} - \widehat{L}_{0.05}}{\widehat{U}_{0.5} - \widehat{L}_{0.5}} - 2.59 \tag{22}$$

where

$$\widehat{U}_\alpha = \frac{1}{n_u(\alpha)} \sum_{y_i > \widehat{P}_{1-\alpha}} y_i, \quad \widehat{L}_\alpha = \frac{1}{n_l(\alpha)} \sum_{y_i < \widehat{P}_\alpha} y_i, \quad \text{for } \alpha \in (0,1)$$

\widehat{P}_α : α^{th} sample percentile

$n_u(\alpha)$: Number of cases satisfying the condition $y_i > \widehat{P}_{1-\alpha}$

$n_l(\alpha)$: Number of cases satisfying the condition $y_i < \widehat{P}_\alpha$

The last measure of kurtosis, which is based on the quantiles was suggested by Crow and Siddiqui [18] and given as follows:

$$\frac{F^{-1}(1 - \alpha) - F^{-1}(\alpha)}{F^{-1}(1 - \beta) - F^{-1}(\beta)} \quad \text{for } \alpha, \beta \in (0,1).$$

Their suggestion for α and β are 0.025 and 0.25, respectively. For standard normal distribution the value of the coefficient was calculated as 2.91 [19] and the centered coefficient was given by

$$kr_4 = \frac{F^{-1}(0.975) - F^{-1}(0.025)}{F^{-1}(0.75) - F^{-1}(0.25)} - 2.91 \tag{23}$$

The following estimator is used for kr_4 .

$$\widehat{kr}_4 = \frac{\widehat{P}_{0.975} - \widehat{P}_{0.025}}{\widehat{P}_{0.75} - \widehat{P}_{0.25}} - 2.91 \tag{24}$$

where \widehat{P}_α is the α^{th} sample percentile.

4. SIMULATION STUDY

In the simulation study, one sided tolerance limit for PCI is used and C_{pu} is computed as a comparison measure. In Rivera et. al.[27], USLs of the non-normal distributions are determined in such a way that fraction of non-conforming parts will give the same fraction of non-conforming parts under normality that corresponds to target PCI value. In this study, similar approach is used to compute USL corresponding to the target C_{pu} . As in Hosseinifard et.al. [14], USL is obtained by equation (25).

$$USL = F^{-1} \left(\Phi(3C_{pu}) \right) \tag{25}$$

where F is the cumulative distribution function of the non-normal data and Φ is the cumulative normal density function.

Target C_{pu} values are taken as 0.5, 1.0, and 1.5. Weibull and lognormal distributions are considered and C_{pu} values are estimated in order to compare with the target values. Estimation of C_{pu} is based on 10000 samples of size $n=100$. For comparative purposes, MSEs are computed and presented in Table 1 and Table 2 when different skewness and kurtosis estimators are used in the implementation of Clements’ method. Table 1 and Table 2 are representing the simulation results for the Weibull distribution and log-

normal distribution, respectively, with different parameter values. In these tables minimum MSEs are given in bold. In order to carry out these simulations a program was written in Matlab.

Table1. MSE of the estimator of C_{pu} for Weibull distribution with different shape parameters when different pairs of skewness and kurtosis estimators are used in the estimation of C_{pu} by Clements’ method.

	C_{pu}		\widehat{kr}_1	\widehat{kr}_2	\widehat{kr}_3	\widehat{kr}_4
wbl(scale=1, shape=3) sk = 0.1681 kr = -0.2705	0.5	\widehat{sk}_1	0.01367	0.00463	0.00560	0.00848
		\widehat{sk}_2	0.01638	0.00481	0.00625	0.00991
		\widehat{sk}_3	0.01430	0.00386	0.00499	0.00859
		\widehat{sk}_4	0.01364	0.00370	0.00480	0.00829
	1	\widehat{sk}_1	0.04777	0.01429	0.01610	0.02839
		\widehat{sk}_2	0.05561	0.01377	0.01897	0.03239
		\widehat{sk}_3	0.04625	0.01115	0.01398	0.02724
		\widehat{sk}_4	0.04526	0.01005	0.01288	0.02658
	1.5	\widehat{sk}_1	0.09074	0.03205	0.03336	0.05688
		\widehat{sk}_2	0.10032	0.02554	0.03218	0.05849
		\widehat{sk}_3	0.08248	0.02121	0.02306	0.04811
		\widehat{sk}_4	0.08212	0.01920	0.02133	0.04516
wbl(scale=1, shape=2) sk = 0.6311 kr = 0.2451	0.5	\widehat{sk}_1	0.01257	0.01094	0.01772	0.03313
		\widehat{sk}_2	0.01565	0.00761	0.00990	0.01463
		\widehat{sk}_3	0.01340	0.00587	0.00761	0.01175
		\widehat{sk}_4	0.01327	0.00591	0.00771	0.01257
	1	\widehat{sk}_1	0.07126	0.17374	0.25984	0.13685
		\widehat{sk}_2	0.11450	0.06544	0.08599	0.11582
		\widehat{sk}_3	0.09667	0.04915	0.06610	0.09368
		\widehat{sk}_4	0.09387	0.05156	0.06968	0.10018
	1.5	\widehat{sk}_1	0.22558	0.42159	0.36326	0.58543
		\widehat{sk}_2	0.35355	0.21913	0.28607	0.37376
		\widehat{sk}_3	0.29608	0.17496	0.22567	0.29965
		\widehat{sk}_4	0.30450	0.18332	0.23408	0.32226
wbl(scale=1, shape=1.5) sk = 1.0720 kr = 1.3904	0.5	\widehat{sk}_1	0.01874	*	*	*
		\widehat{sk}_2	0.01644	0.01003	0.01222	0.01686
		\widehat{sk}_3	0.01211	0.00765	0.00910	0.01345
		\widehat{sk}_4	0.01174	0.00822	0.00982	0.01485
	1	\widehat{sk}_1	0.10922	*	*	*
		\widehat{sk}_2	0.13769	0.19074	0.22086	0.25982
		\widehat{sk}_3	0.12543	0.14987	0.17429	0.22243
		\widehat{sk}_4	0.12849	0.16295	0.18806	0.22984
	1.5	\widehat{sk}_1	0.55984	*	*	*
		\widehat{sk}_2	0.65436	0.91485	1.02458	1.17467
		\widehat{sk}_3	0.56473	0.76900	0.85416	1.01083
		\widehat{sk}_4	0.59813	0.82026	0.90437	1.04805

* : represents very large number

According to Table 1, for small values of skewness and kurtosis, using traditional estimators (\widehat{sk}_1 - \widehat{kr}_1) in the estimation of C_{pu} results in larger MSE as compared to the use of other estimators of skewness and kurtosis. Especially, using \widehat{sk}_3 or \widehat{sk}_4 in combination with \widehat{kr}_2 in the estimation of C_{pu} by Clements’ method might be preferred by the practitioners since their corresponding MSE is considerably less than the MSE of \widehat{C}_{pu} which is obtained by traditional estimators.

Table2. MSE of the estimator of C_{pu} for lognormal distribution with different shape parameters when different pairs of skewness and kurtosis estimators are used in the estimation of C_{pu} by Clements’ method.

	Cpu		\widehat{kr}_1	\widehat{kr}_2	\widehat{kr}_3	\widehat{kr}_4
$\log n(\mu = 0, \sigma = 0.05)$ $sk = 0.1502 \quad kr = 0.0401$	0.5	\widehat{sk}_1	0.01147	0.00657	0.00696	0.00922
		\widehat{sk}_2	0.01008	0.00458	0.00491	0.00761
		\widehat{sk}_3	0.00869	0.00400	0.00437	0.00651
		\widehat{sk}_4	0.00825	0.00376	0.00412	0.00639
	1	\widehat{sk}_1	0.05194	0.02982	0.02930	0.04254
		\widehat{sk}_2	0.04835	0.02234	0.02565	0.03728
		\widehat{sk}_3	0.04316	0.01758	0.02007	0.03042
		\widehat{sk}_4	0.04161	0.01675	0.01959	0.03072
	1.5	\widehat{sk}_1	0.15072	0.07758	0.09156	0.11947
		\widehat{sk}_2	0.15308	0.07619	0.08393	0.11682
		\widehat{sk}_3	0.12955	0.06264	0.06938	0.09911
		\widehat{sk}_4	0.12515	0.05946	0.06873	0.09347
$\log n(\mu = 0, \sigma = 0.10)$ $sk = 0.3018 \quad kr = 0.1623$	0.5	\widehat{sk}_1	0.01176	0.00609	0.00711	0.00878
		\widehat{sk}_2	0.01034	0.00510	0.00582	0.00834
		\widehat{sk}_3	0.00949	0.00455	0.00487	0.00705
		\widehat{sk}_4	0.00928	0.00441	0.00471	0.00720
	1	\widehat{sk}_1	0.05992	0.15722	0.03642	0.16736
		\widehat{sk}_2	0.07135	0.04122	0.04437	0.05788
		\widehat{sk}_3	0.06320	0.03246	0.03645	0.04850
		\widehat{sk}_4	0.06066	0.03203	0.03637	0.04821
	1.5	\widehat{sk}_1	0.22658	0.35713	0.14682	0.29173
		\widehat{sk}_2	0.28444	0.18628	0.20959	0.24481
		\widehat{sk}_3	0.25196	0.16206	0.17658	0.20935
		\widehat{sk}_4	0.25053	0.15857	0.17863	0.21191
$\log n(\mu = 0, \sigma = 0.40)$ $sk = 1.3219 \quad kr = 3.26$	0.5	\widehat{sk}_1	0.02719	*	*	*
		\widehat{sk}_2	0.01466	0.00984	0.01022	0.01201
		\widehat{sk}_3	0.02935	0.00847	0.00927	0.01155
		\widehat{sk}_4	0.01634	0.00844	0.00949	0.01132
	1	\widehat{sk}_1	0.17716	*	*	*
		\widehat{sk}_2	0.28336	0.42216	0.41290	0.41566
		\widehat{sk}_3	0.27426	0.36440	0.34663	0.35257
		\widehat{sk}_4	0.22575	0.38488	0.36941	0.37907
	1.5	\widehat{sk}_1	2.37022	*	*	*
		\widehat{sk}_2	3.29255	4.60860	4.44679	4.40649
		\widehat{sk}_3	2.75485	4.08871	3.97120	3.95354
		\widehat{sk}_4	3.98982	4.29578	4.16337	4.09141

* : represents very large number

As the skewness and kurtosis of the process data increase, using traditional estimators of them results in smaller MSE of \widehat{C}_{pu} . Unless C_{pu} is very small, it is better to use traditional estimators ($\widehat{sk}_1 - \widehat{kr}_1$) in the estimation of C_{pu} .

For large values of skewness and kurtosis, the use of \widehat{kr}_2 , \widehat{kr}_3 and \widehat{kr}_4 with \widehat{sk}_1 gives very large MSE values of \widehat{C}_{pu} . Therefore, $\widehat{sk}_1 - \widehat{kr}_2$, $\widehat{sk}_1 - \widehat{kr}_3$ and $\widehat{sk}_1 - \widehat{kr}_4$ pairs should not be preferred in the estimation of C_{pu} .

When we examine Table 2, for small values of skewness and kurtosis, generally using \widehat{sk}_4 and \widehat{kr}_2 pair in the estimation of C_{pu} by Clement's method results in smaller MSE as compared to the use of other estimators of skewness and kurtosis. As skewness and kurtosis of the data get larger MSE values become very large, especially when $\widehat{sk}_1 - \widehat{kr}_2$, $\widehat{sk}_1 - \widehat{kr}_3$, and $\widehat{sk}_1 - \widehat{kr}_4$ pairs are used in the estimation of C_{pu} . Therefore, if the data has large skewness and kurtosis, instead of these pairs, traditional estimators of

skewness and kurtosis, \widehat{sk}_1 and \widehat{kr}_1 , should be preferred for estimation of C_{pu} . However, if the skewness and kurtosis of the data are not so large one can prefer to use $\widehat{sk}_4 - \widehat{kr}_2$ pair in the estimation of C_{pu} by Clement’s method.

5. A REAL LIFE APPLICATION TO OIL PUMP MANUFACTURING

In an engine oil pump, there are several important quality characteristics that have to conform to product specifications determined by the design engineers so that the oil pump performs as expected. It is quite significant that those quality characteristics are individually monitored to their tolerance limits and effects of tolerance stack-up due to those limits, which can simply be expressed as the accumulation of dimensional errors, are kept under control. Highly accurate and precise CMMs (Coordinate Measuring Machine) are used to measure those quality characteristics knowing that any dimensional non-conformity may negatively influence product performance or even lead to total product failure causing drastic car engine break. Measurement results are then used to perform process capability inspections to ensure that the process is under control. A correct evaluation of process capability requires proper analysis of the available data, determination of the type of statistical distribution and then computation of the process capability indices. Distributions of the most of the quality characteristics in oil pump manufacturing process do not fit to normal distribution contrary to expectations. Therefore, standard PCI’s may cause misleading interpretations.

This study aims investigation and interpretation of machining process capability for body part pocket bottom surface flatness which has considerable effect on product performance. Any undetected non-conformity of this characteristic has major potential consequences such as insufficient engine cooling and thus low engine performance, rapid wear on engine components thus shorter engine life and in the worst case, complete engine burn out during end customer use.

In the present application, tolerance limit of the pocket bottom surface flatness was defined as 30 microns by the technical drawing. Measurements were carried out on one single CMM. Parts were located on fixtures using three support surfaces of the body part. After measurement datums were set by CMM, 1500 different coordinate measurements were made for each part from pre-defined scattered positions representing the whole surface as a point cloud. By the help of those measurements, a virtual flat surface was fit by CMM software using Least-Squares Fitting Algorithms based on orthogonal distance regression. The orthogonal distance between the upper and lower peak points of the point cloud along the normal of the fit flat surface gave the flatness value.

Totally 233 pcs were inspected in this study in order to correctly calculate the capability indices and thus make proper interpretation of the whole flatness data.

Initially the histogram and normal probability plot for the flatness data are obtained and given in Figure 1.

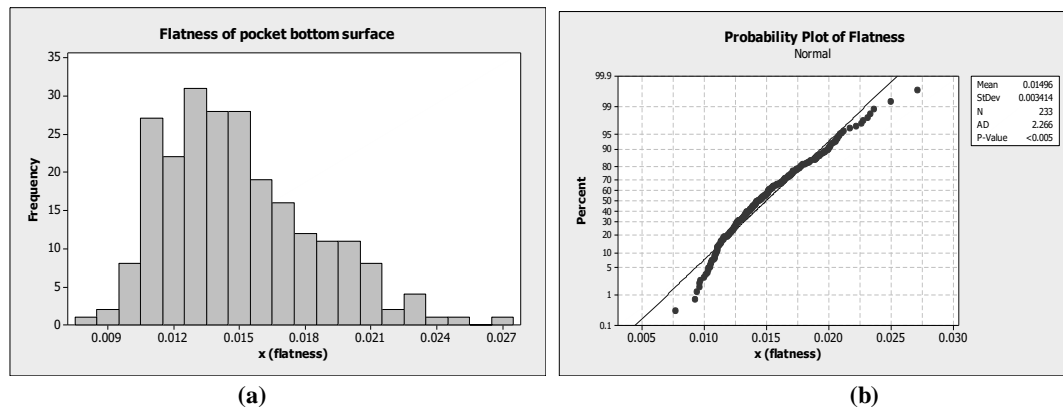


Figure 1. (a) Histogram of pocket bottom flatness data (b) Normal probability plot of pocket bottom flatness data

When the histogram and the normal probability plot are examined one can easily conclude that the flatness data is not distributed as Normal (p value < 0.005). Therefore, it is not appropriate to use standard PCIs in the capability analysis and the application of Clements’ method with the possible different estimators of skewness and kurtosis given in section 3 is presented in this section.

The procedure for calculating PCIs by Clements’ method given in section 2 is conducted for the pocket bottom flatness data.

1. Note that the LSL and USL were 0.00 and 0.03 mm, respectively. However, since flatness can not take a negative value, this process characteristic can be considered as having just USL = 0.03 mm.
2. The process mean and the standard deviation are estimated.
 $\bar{x} = 0.014962, s = 0.003414$
3. Skewness and kurtosis of the flatness data are estimated by both traditional estimators and by the other robust estimators given in section 3. The obtained estimates are given in Table 3.

Table 3. Skewness and kurtosis estimates of pocket bottom flatness data.

\widehat{sk}_1	0.6832	\widehat{kr}_1	0.1794
\widehat{sk}_2	0.1460	\widehat{kr}_2	0.0903
\widehat{sk}_3	0.2096	\widehat{kr}_3	-0.1704
\widehat{sk}_4	0.1663	\widehat{kr}_4	-0.1314

4. Standardized values of L_p, U_p and M (denoted by L'_p, U'_p and M') are obtained from the tables in [6] and given in Table 4.

Table 4. Standardized values of L_p, U_p and M obtained for different estimates of skewness and kurtosis of pocket bottom flatness data.

		$\widehat{kr}_1 = 0.1794$	$\widehat{kr}_2 = 0.0903$	$\widehat{kr}_3 = -0.1704$	$\widehat{kr}_4 = -0.1314$
$\widehat{sk}_1 = 0.6832$	L'_p	1.7526	1.6858	1.4929	1.5216
	U'_p	3.5287	3.4408	3.1541	3.1993
	M'	0.1689	0.1620	0.1790	0.1831
$\widehat{sk}_2 = 0.1460$	L'_p	2.9020	2.8410	2.6383	2.6705
	U'_p	3.2941	3.2322	3.0307	3.0624
	M'	0.0266	0.0257	0.0270	0.0274
$\widehat{sk}_3 = 0.2096$	L'_p	2.8016	2.7367	2.5282	2.5609
	U'_p	3.3575	3.2950	3.0914	3.1235
	M'	0.0378	0.0368	0.0395	0.0401
$\widehat{sk}_4 = 0.1663$	L'_p	2.8708	2.8085	2.6039	2.6362
	U'_p	3.3149	3.2529	3.0508	3.0827
	M'	0.0301	0.0292	0.0309	0.0314

5. L_p, U_p and M are estimated as described in section 2 and they are given in Table 5.

Table 5. Calculated values of L_p , U_p and M for pocket bottom flatness data.

		$\widehat{kr}_1 = 0.1794$	$\widehat{kr}_2 = 0.0903$	$\widehat{kr}_3 = -0.1704$	$\widehat{kr}_4 = -0.1314$
$\widehat{sk}_1 = 0.6832$	L_p	0.0090	0.0092	0.0099	0.0098
	U_p	0.0270	0.0267	0.0257	0.0259
	M	0.0155	0.0155	0.0156	0.0156
$\widehat{sk}_2 = 0.1460$	L_p	0.0051	0.0053	0.0060	0.0058
	U_p	0.0262	0.0260	0.0253	0.0254
	M	0.0151	0.0150	0.0151	0.0151
$\widehat{sk}_3 = 0.2096$	L_p	0.0054	0.0056	0.0063	0.0062
	U_p	0.0264	0.0262	0.0255	0.0256
	M	0.0151	0.0151	0.0151	0.0151
$\widehat{sk}_4 = 0.1663$	L_p	0.0052	0.0054	0.0061	0.0060
	U_p	0.0263	0.0261	0.0254	0.0255
	M	0.0151	0.0151	0.0151	0.0151

6. For the flatness data there is just USL given by the design engineers. Therefore, C_{pu} values are estimated by using equation (4) in order to assess the capability of the process and they are given in Table 6.

Table 6. Estimated C_{pu} values for pocket bottom flatness data.

	$\widehat{kr}_1 = 0.1794$	$\widehat{kr}_2 = 0.0903$	$\widehat{kr}_3 = -0.1704$	$\widehat{kr}_4 = -0.1314$
$\widehat{sk}_1 = 0.6832$	1.26	1.29	1.42	1.40
$\widehat{sk}_2 = 0.1460$	1.34	1.37	1.46	1.44
$\widehat{sk}_3 = 0.2096$	1.32	1.34	1.43	1.42
$\widehat{sk}_4 = 0.1663$	1.33	1.36	1.45	1.43

In Table 6, C_{pu} is estimated by using all possible estimators of skewness and kurtosis given in this study. When the histogram and the estimates of skewness and kurtosis of pocket bottom flatness data are examined, we see that the skewness of the flatness data is not so large and the kurtosis estimates fluctuate around 0. In the previous section, simulation results showed us that if the skewness and kurtosis are not so large, instead of using traditional estimators of them it is better to use $\widehat{sk}_3 - \widehat{kr}_2$ or $\widehat{sk}_4 - \widehat{kr}_2$ pair in the estimation of C_{pu} . Therefore, C_{pu} values that are estimated by using these pairs are given in bold in Table 6. Assessing the process capability with the C_{pu} values estimated by the use of the traditional estimators ($\widehat{sk}_1 - \widehat{kr}_1$) may mislead the practitioner about the capability of the process. C_{pu} for the flatness data which is estimated by the traditional estimators of skewness and kurtosis is smaller than the C_{pu} estimated by using $\widehat{sk}_3 - \widehat{kr}_2$ or $\widehat{sk}_4 - \widehat{kr}_2$ pair, and smaller C_{pu} is the indication of less capable process. In that case a practitioner may take an action although it is not actually needed.

6. CONCLUSION

When probability distribution of a process characteristic is non-normal, PCIs calculated using the traditional methods often causes misleading interpretation of process capability. Several methods were proposed in the literature in order to analyze the capability of process when the process characteristic is non-normal. One of the most widely discussed methods to handle non-normality is Clements' method. In this study, Clements' method was modified by changing the estimators of skewness and kurtosis used in the estimation procedure of PCIs. Instead of using traditional estimators of skewness and kurtosis, more robust estimators proposed in the literature were used in the estimation of PCIs by Clements' method. For comparison purposes, simulation study was conducted and results were obtained when a process characteristic has Weibull or log-normal distribution. Simulation results showed that if the

distribution of the process characteristic is not so much skewed and its kurtosis is not large then instead of using the traditional estimators of skewness and kurtosis it is better to use $\widehat{sk}_3 - \widehat{kr}_2$ or $\widehat{sk}_4 - \widehat{kr}_2$ pair in the estimation of C_{pu} by Clements' method. A real life example regarding an oil pump manufacturing process was provided in section 5, and pocket bottom flatness characteristic in the engine oil pump was analyzed by estimating the C_{pu} index in order to make interpretations about the capability of the process. Several C_{pu} values were estimated by using different estimators of skewness and kurtosis. However, the use of \widehat{C}_{pu} which utilizes $\widehat{sk}_3 - \widehat{kr}_2$ or $\widehat{sk}_4 - \widehat{kr}_2$ pair in the estimation procedure was suggested relying on the simulation results presented in section 4.

REFERENCES

- [1] Box GEP, Cox DR. An analysis of transformations. *J Roy Statist Soc B* 1964; 26: 211-243.
- [2] Johnson NL, Kotz S. *Discrete Distributions*. New York, NY, USA: John Wiley, 1969.
- [3] Ryan TP. *Statistical Methods for Quality Improvement*. New York, NY, USA: John Wiley, 1989.
- [4] Ryan TP, Schwertman NC. Optimal limits for attributes control charts. *J Qual Technol* 1997; 29: 86-96.
- [5] Niaki STA, Abbasi B. Skewness reduction approach in multi-attribute process monitoring. *Commun Stat - Theory Methods* 2007; 36:12, 2313-2325.
- [6] Clements JA. Process capability indices for non-normal calculations. *Qual Prog* 1989; 22: 49-55.
- [7] Liu PH, Chen FL. Process capability analysis of non-normal process data using the Burr XII distribution. *Int J Adv Manuf Technol* 2006; 27: 975-984.
- [8] Wu HH, Swain JJ, Farrington PA, Messimer SH. A weighted variance capability index for general non-normal processes. *Qual Reliab Engng Int* 1999; 15: 397-402.
- [9] Choobineh F, Branting D. A simple approximation for semivariance. *Eur J Oper Res* 1986; 27: 364-370.
- [10] Chang YS, Choi IS, Bai DS. Process capability indices for skewed populations. *Qual Reliab Engng Int* 2002; 18: 383-393.
- [11] Tang LC, Than SE. Computing process capability indices for non-normal data: a review and comparative study. *Qual Reliab Engng Int* 1999; 15: 339-353.
- [12] Kovarik M, Sarga L. Process capability indices for non-normal data. *Wseas Transactions on Business and Economics* 2014; 11: 419-429.
- [13] Sennaroglu B, Senvar O. Performance comparison of Box-Cox transformation and weighted variance methods with Weibull distribution. *Journal of Aeronautics and Space Technologies* 2015; 8: 49-55.
- [14] Hosseinifard SZ, Abbasi B, Ahmad S, Abdollahian M. A transformation technique to estimate the process capability index for non-normal processes. *Int J Adv Manuf Technol* 2009; 40: 512-517.
- [15] Kotz S, Johnson NL. Process capability indices-a review, 1992-2000 (with discussions). *J Qual Technol* 2002; 34: 2-53.

- [16] Anis MZ. Basic process capability indices: an expository review. *Int Stat Rev* 2008; 76:3, 347-367.
- [17] Wu CW, Pearn WL, Kotz S. An overview of theory and practice on process capability indices for quality assurance. *Int J Prod Econ* 2009; 117: 338–359.
- [18] Crow EL, M.M. Siddiqui MM. Robust estimation of location. *J Am Stat Assoc* 1967, 62: 353–389.
- [19] Kim T, White H. On more robust estimation of skewness and kurtosis. *Financ Res Lett* 2004; 1: 65–70.
- [20] Bowley AL. *Elements of Statistics*. Scribner's, New York, 1920.
- [21] Hinkley DV. On power transformations to symmetry. *Biometrika* 1975; 62: 101–111.
- [22] Groeneveld RA, Meeden G. Measuring skewness and kurtosis. *The Statistician* 1984; 33: 391–399.
- [23] Kendall MG, Stuart A. *The Advanced Theory of Statistics*. Griffin, London, 1977.
- [24] Moors JJA. A quantile alternative for kurtosis. *The Statistician* 1988; 37: 25–32.
- [25] Hogg RV. More light on the kurtosis and related statistics. *J Am Stat Assoc* 1972, 67: 422-424.
- [26] Hogg RV. Adaptive robust procedures: A partial review and some suggestions for future applications and theory. *J Am Stat Assoc* 1974, 69: 909–923.
- [27] Rivera LAR, Hubele NF, Lawrence FP. C_{pk} index estimation using data transformation. *Comput Ind Eng* 1995, 29: 55:58.