Mazandaran University of Science and Technology

Available online at www.iiec2017.com

13th International Conference on Industrial Engineering (IIEC 2017)

Iran Institute of Industrial Engineering

Robust Estimation of Process Capability Indices

Behrooz Khalilloo[a] , Hamid Shahriari[b] and Emad Roghanian[c]

[a] *Graduate Student, Department of Industrial Engineering*
K. N. Toosi University of Technology, Tehran, Iran
E-mail: behrooz.khalilloo@gmail.com

[b] *Professor, Department of Industrial Engineering*
K. N. Toosi University of Technology, Tehran, Iran
E-mail: hshahriari@kntu.ac.ir

[c] *Associate Professor, Department of Industrial Engineering*
K. N. Toosi University of Technology, Tehran, Iran
E-mail: e_roghanian@kntu.ac.ir

Abstract

The initial step in statistical analysis is parameter estimation. In univariate analysis, the parameters mean and standard deviation must be estimated when they are unknown. when outliers exist in data, use of sample mean results in week estimation. So, estimators which are robust to the presence of outliers should be used. In this work robust M-estimator for estimating those parameters are used. The performance of these robust estimators in presence of outliers and their effects on process capability indices are studied. The results indicate that the proposed robust capability indices perform much better than the existing process capability indices.

Keywords:
Process Capability Indices, Robust Estimator, M-Estimator, Univariate

Introduction

Process capability indices have been widely used in industries to evaluate the performance of the processes. Various different univariate capability indices have been introduced. In univariate analysis, it is usually assumed that the data come from a normal distribution. In this situation, the maximum likelihood estimator (MLE), is good estimator. However, Montgomery in [3] stated that when the process is not in control, its parameters are unstable. Therefore, an important consideration for determining the process capability is to know whether the process is in statistical control or not. Robust estimator such as M-estimator and other estimators are introduced in [1,2].
The first process capability index introduced in [4] is Cp that measures the potential capability of a process with no attention to the process mean. It is defined as:

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

where USL and LSL are the upper and the lower specification limits for quality characteristic, respectively and σ is the process standard deviation. Since process standard deviation is unknown, it must be estimate, so the estimate of Cp would be:

$$\hat{C}p = \frac{USL - LSL}{6\hat{\sigma}} \quad (2)$$

where $\hat{\sigma}$ is the estimate of process standard deviation.
The second process capability index introduced in [5]. The Cpk index is a measure of the capability of a process with attention to the process mean.

$$Cpk = \min\{Cpu, Cpl\} \quad (3)$$

$$Cpu = \frac{USL - \mu}{3\sigma} \quad (4)$$

$$Cpl = \frac{\mu - LSL}{3\sigma} \quad (5)$$

where μ is the process mean.
The third process capability is Cpm introduced in [6] is defined as:

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

where T is the process target value.
The Cpkm is another one which is defined as:

$$Cpmk = \frac{\min(USL - \mu, \mu - LSL)}{3\sqrt{\sigma^2 + (\mu - T)^2}} \quad (7)$$

In reference [7] a process capability is proposed which unifies the four basic process capability indices (PCIs), Cp, Cpk, Cpm and Cpmk as follows:

$$Cp(u,v) = \frac{d - u|\mu - m|}{3\sqrt{\sigma^2 + v(\mu - T)^2}} \quad (8)$$

where u and v are constants that take value 0 or 1 and T is the target value. When T is unknown, then T = m is often used [8], where m is the center of the specification limits, μ is the process mean and σ is the process standard deviation.
We can Achieve four basic univariate process capability indices by:

$$C_p = C_p(0,0),\quad C_{pk} = Cp(1,0),\quad C_{pm} = Cp(0,1),\quad C_{pkm} = Cp(1,1) \quad (9)$$

Methodology

Suppose we have a sample from population and this sample has some outliers. When we estimate the mean and the standard deviation of the process by MLE or other methods, our estimators which will be used to estimate the process capability indices, certainly make bad estimation for process capability indices that will confuse the manager to make a good decision for improving the performance of the process.
In this work we estimate the mean and the standard deviation of the process with the robust M-estimator and to eliminate the effect of the outlier data on the estimation of the mean and the standard deviation of the process.

M-estimator

One of the most popular robust estimators is M-estimator. Robust M-estimators are the modified MLEs, and according to [1] are less sensitive to the presence of outlying observations. Based on a random sample $\{x_1, x_2, ..., xn\}$ the M-estimators of μ and σ are as:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} w_{1i}(x_i - \mu) * x_i}{\sum_{i=1}^{n} w_{1i}(x_i - \mu)} \quad (10)$$

$$\hat{\sigma}^2 = \frac{1}{n\delta}\sum_{i=1}^{n} w_{2i}(x_i/\sigma)x_i^2 \quad (11)$$

where
$w_1(x) = \psi(x)/x$, $w_2 = \rho(x)/x^2$, $\psi(x) = d\rho(x)/dx$ and $\delta = 0.5$.

Based on [2], choosing some suitable $\rho(x)$ function would yield robust estimators. It must be noted that for an M-estimator, the $w_1(x)$, $w_2(x)$ functions in Equations (10) and (11) should not be the same. w_1 is a weight function defined as $w(x) = \min\{1, k/|x|\}$, with k= 1.37.A frequently used scale estimate is the bisquare scale that for the bisquare scale $w_2(x) = \min(1، (1-(x/k)^2)^2)$ with k=4.68.

Examples

In this section we will show that for a sample that may have some outliers, the robust estimators have a good estimate for process mean and standard deviation. In result we can estimate process capability indices more accurately and manager can make better decision about the process improvement and to take some actions for process improvement.

Example 1

In this example, piston ring inside diameter from [3] is studied. The data of piston ring inside diameter are collected with USL=74.005, LSL=73.995 with T=74. The for this example is in Table 1 and plot of the data is shown in Figure 1 and We want to find the robust estimates of the mean and the standard deviation of piston ring inside diameter from the sample data. As mentioned in [3],the estimated values are:

$\bar{\bar{X}} = 74.001$ *and* $\hat{\sigma} = 0.01$.

Now with proposed M-estimator introduced in section 2, μ and σ are estimated as:

$$\hat{\mu} = 74.001 \;\; and \;\; \hat{\sigma} = 0.0091$$

It is clear that the mean estimated as a robust estimator don't change but the standard deviation is improved. By calculation of four basic univariate PCIs from traditional methods and robust M-estimator and comparison of the results we see that the robust estimator has a good estimate for the process mean and the standard deviation. Results for PCIs are summarized in Table 2 and comparison of the two method is shown in Figure 2.

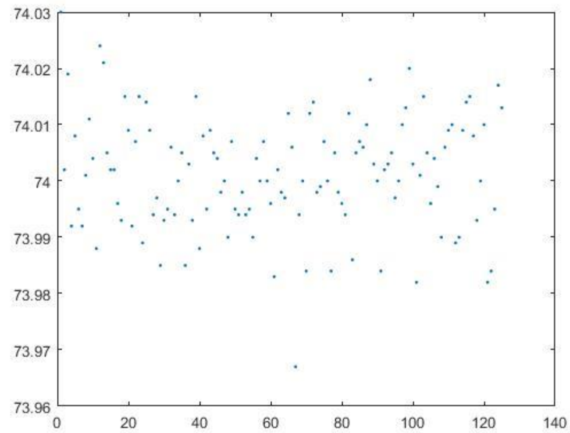

Figure 1 – Plot of piston ring inside diameter Data

Table 1 - piston ring inside diameter Data

74.03	74.002	74.019	73.992	74.008
73.995	73.992	74.001	74.011	74.004
73.988	74.024	74.021	74.005	74.002
74.002	73.996	73.993	74.015	74.009
73.992	74.007	74.015	73.989	74.014
74.009	73.994	73.997	73.985	73.993
73.995	74.006	73.994	74	74.005
73.985	74.003	73.993	74.015	73.988
74.008	73.995	74.009	74.005	74.004
73.998	74	73.99	74.007	73.995
73.994	73.998	73.994	73.995	73.99
74.004	74	74.007	74	73.996
73.983	74.002	73.998	73.997	74.012
74.006	73.967	73.994	74	73.984
74.012	74.014	73.998	73.999	74.007
74	73.984	74.005	73.998	73.996
73.994	74.012	73.986	74.005	74.007
74.006	74.01	74.018	74.003	74
73.984	74.002	74.003	74.005	73.997
74	74.01	74.013	74.02	74.003
73.982	74.001	74.015	74.005	73.996
74.004	73.999	73.99	74.006	74.009
74.01	73.989	73.99	74.009	74.014
74.015	74.008	73.993	74	74.01
73.982	73.984	73.995	74.017	74.013

Table 2 – PCIs value for usual and robust method for piston ring inside diameter data

Index Method	Cp	Cpm	Cpk	Cpkm
MLE	0.1667	0.1333	0.1658	0.1327
Robust M-Estimator	0.1852	0.1481	0.1841	0.1472

Example 2

In this example, a hard-bake process from [3] is studied. The data of hard-bake process are with USL=2, LSL=1 and T=1.5. The ddata for this example is in Table 3 and plot of the data is shown in Figure 3. We want to find the robust estimate of the mean and the standard deviation of the hard-bake process from the sample data. As mentioned in [3], estimated values are:

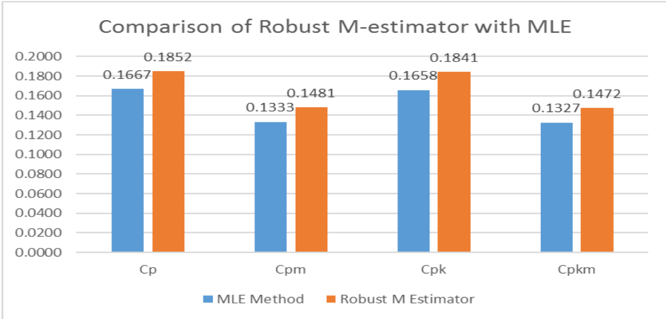

Figure 2 - Comparison of Robust M Estimator with Usual Method for piston

$$\bar{\bar{X}} = 1.5056, \hat{\sigma} = 0.1398$$

Now with proposed M-estimator introduced in section 2, μ and σ are estimated as:

$$\hat{\mu} = 1.5075, \hat{\sigma} = 0.1267 \; .$$

It is clear that the robust mean and the robust standard deviation estimates are improved. By calculation of four basic univariate PCIs from traditional methods and robust M-estimator and comparison of the results we see that the robust estimators of process capability indices are more realistic and all represent same values. Results for PCIs are summarized in Table 4 and comparison of the two methods is shown in Figure 4.

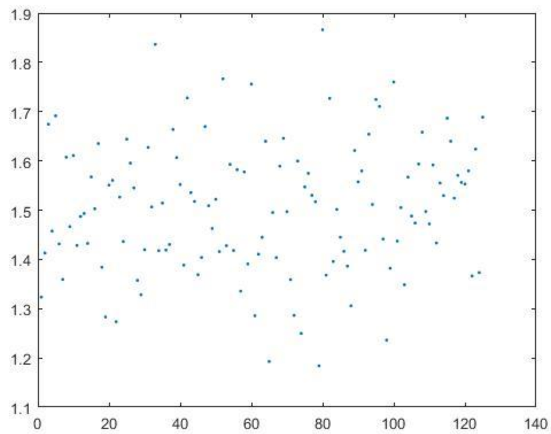

Figure 3 – Plot of Hard-bake Process Data

Table 3 - Hard-bake Process Data

1.3235	1.4128	1.6744	1.4573	1.6914
1.4314	1.3592	1.6075	1.4666	1.6109
1.4284	1.4871	1.4932	1.4324	1.5674
1.5028	1.6352	1.3841	1.2831	1.5507
1.5604	1.2735	1.5265	1.4363	1.6441
1.5955	1.5451	1.3574	1.3281	1.4198
1.6274	1.5064	1.8366	1.4177	1.5144
1.419	1.4303	1.6637	1.6067	1.5519
1.3884	1.7277	1.5355	1.5176	1.3688
1.4039	1.6697	1.5089	1.4627	1.522
1.4158	1.7667	1.4278	1.5928	1.4181
1.5821	1.3355	1.5777	1.3908	1.7559
1.2856	1.4106	1.4447	1.6398	1.1928
1.4951	1.4036	1.5893	1.6458	1.4969
1.3589	1.2863	1.5996	1.2497	1.5471
1.5747	1.5301	1.5171	1.1839	1.8662
1.368	1.7269	1.3957	1.5014	1.4449
1.4163	1.3864	1.3057	1.621	1.5573
1.5796	1.4185	1.6541	1.5116	1.7247
1.7106	1.4412	1.2361	1.382	1.7601
1.4371	1.5051	1.3485	1.567	1.488
1.4738	1.5936	1.6583	1.4973	1.472
1.5917	1.4333	1.5551	1.5295	1.6866
1.6399	1.5243	1.5705	1.5563	1.553
1.5797	1.3663	1.624	1.3732	1.6887

Table 4 - PCIs value for usual and robust method for Hard-bake Process Data

Index Method	Cp	Cpm	Cpk	Cpkm
MLE Method	1.1922	1.1788	1.1912	1.1779
Robust M Estimator	1.3154	1.3004	1.3141	1.2991

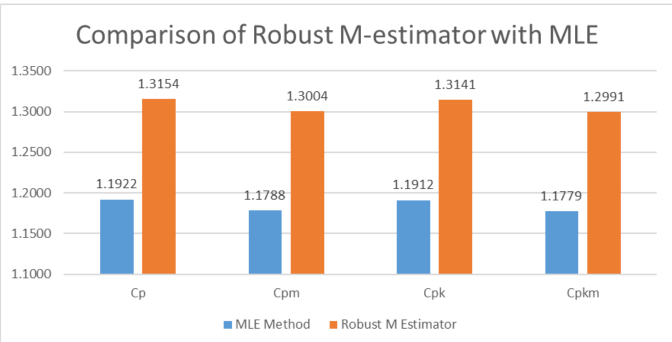

Figure 4 - Comparison of Robust M Estimator with Usual Method for Hard-bake Process Data

Example 3

In this example, the layer thickness from [3] is studied. The data of layer thickness are with USL=480, LSL=420 and T=450. The data for this example data is in Table 5 and plot of data is shown in Figure 5. We want to find the robust estimate of the mean and the standard deviation of the Layer Thickness from the sample data. As mentioned in [3], estimated values are:

$$\bar{\bar{X}} = 448.687, \hat{\sigma} = 8.0136$$

Now with proposed M-estimator introduced in section 2, μ and σ are estimated as:

$$\bar{\bar{X}} = 448.9, \hat{\sigma} = 7.2647$$

It is clear that the mean and the standard deviation estimated as a robust estimators are improved. By calculation of four basic univariate PCIs from traditional methods and the robust M-estimator and comparison of the results we see that the robust estimator estimators of process capability indices are more realistic and all represent same values. Results for PCIs are summarized in Table 6 and comparison of the two method is shown in Figure 6.

Table 5 - Layer Thickness Data

459	449	435	450
443	440	442	442
457	444	449	444
469	463	453	438
443	457	445	454
444	456	456	457
445	449	450	445
446	455	449	452
444	452	457	440
432	463	463	443
445	452	453	438
456	457	436	457
459	445	441	447
441	465	438	450
460	453	457	438
453	444	451	435
451	460	450	457
422	431	437	429
444	446	448	467
450	450	454	454

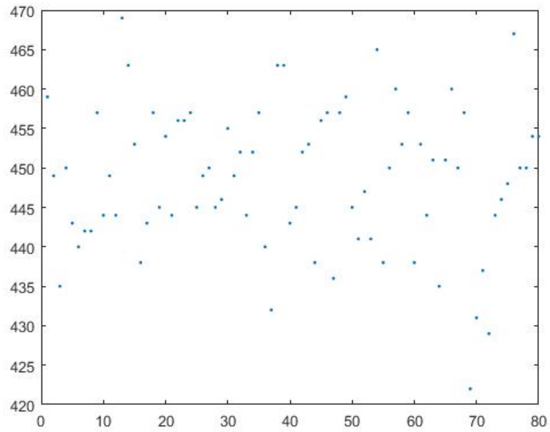

Figure 5 – Plot of Layer Thickness Data

Table 6 - PCIs value for usual and robust method for Layer Thickness Data

Index Method	Cp	Cpm	Cpk	Cpkm
MLE Method	1.2479	1.1933	1.2315	1.1776
Robust M Estimator	1.3765	1.3077	1.3480	1.2806

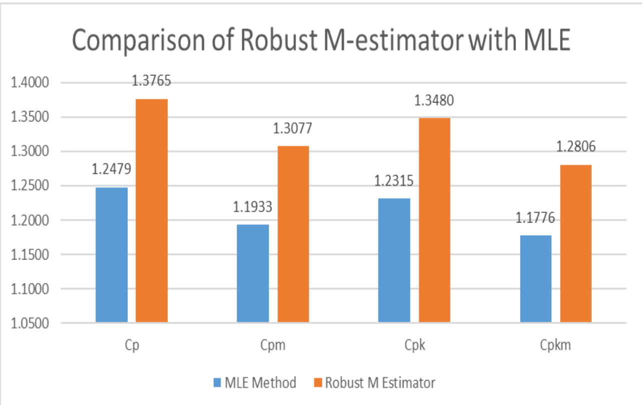

Figure 6 - Comparison of Robust M Estimator with Usual Method for Layer Thickness Data

Discussion and conclusion

It is shown by examples that the robust m-estimator has better estimate than the MLE method. In robust method if a data has big distant from process mean, by using weight function, all of the data do not have a same weight and their weight depend on their distant from the process mean unlike MLE method that all data have same weight. In real operation, when the manager decides to make necessary changes to improve the process, it is important to have true information about the process that when we have outlier(s) in our data, the robust estimates have better performance. Use of the robust estimators of the process parameter in defining the process capability indices in presence of outliers result a true value for the index under consideration.

References

[1] P.J. Huber and E.M. Ronchetti, *Robust Statistics*, 2nd ed., Wiley and Sons, Hoboken, NJ, 2009

[2] R.A. Maronna, R.D. Martin, and V.J. Yohai, *Robust Statistics, Theory and Methods*, Wiley and Sons, New York,
2006

[3] Montgomery DC (2009) *Introduction to statistical quality control*. Wiley, New York

[4] Kane, V. E. (1986a). Process Capability Indices. *J. of Quality Technology* 18, 41-52

[5] Sullivan, L.P. (1985). Letters. *Quality Progress* 18, 7 - 8

[6] Chan, L. K., Cheng, S. W., and Spiring, F. A. (1988). A new Measure of Process Capability: Cpm. *Journal of Quality Technology* 20,162 - 175.

[7] Vännman, K. (1995). A Unified Approach to Process Capability Indices. *Statistical Sinica* 5, 805 - 820

[8] Pearn,W.L. and Kotz, S. (2006). Encyclopedia and Handbook of Process Capability Indices: A comprehensive exposition of quality control measures, World Scientific Publishing Co. Ltd, Singapore.