Derivation of the Limits for Control Chart Using the Median Absolute Deviation for Monitoring Non-Normal Process

Adekeye, K.S. and P.I. Azubuike
Department of Mathematical Sciences,
College of Natural Sciences, Redeemers University (RUN), Nigeria

Abstract: Problem statement: The Shewhart and S control charts, in the literature, were combined to evaluate the stability of a process. These charts were based on the fundamental assumption of normality of the quality characteristics under investigation. **Approach:** In practice, the normality assumption was often violated by real life data, therefore, use of the Shewhart and S control charts on real life data might leads to misplacement of control limits. There were many alternatives in the literature to handle non-normality of quality characteristics. The Median Absolute Deviation (MAD) claimed in the literature to be the best estimate when the data under consideration is non-normal. Thus in this study, we derived the control limits for the-control chart using the median absolute deviation for monitoring process stability when the quality characteristic under investigation was non-normal. **Results:** The derived control limits were compared with the control limits when the sample standard deviation was used as a measure of controlling the process variability using manufacturing process (real life) data. Furthermore, a simulation study was carried out to evaluate the performance of the proposed MAD based control charts on both normal and non-normal process. **Conclusion:** The obtained results show that the derived control limit is an improvement on the control limit of the Shewhart and that the MAD control charts performed better for non-normal process than for normal process.

Key words: Median absolute deviation, non-normal, control limits interval, process variability, statistical process, control charts, sigma approach, standard deviation, soft drink, Cofta tablet, manufacturing process

INTRODUCTION

The Shewhart $\bar{\mathrm{X}}$-control chart is one of the most widely used statistical process control techniques developed to monitor/control the process average while the S- chart is used to monitor the process variability. The two charts called $\bar{\mathrm{X}}$ and S control charts in the literature, are combined to evaluate the stability of a process. In $\bar{\mathrm{X}}$ control chart, if the average of a subgroup falls beyond the control limits, it is an indication that the process average is out of control and in S control chart, if the standard deviation of a subgroup falls beyond the control limits, it is an indication that the process variability is out of control (Adekeye, 2009; Montgomery, 2008). However, these charts are based on the fundamental assumption of normality of the quality characteristics under investigation. In the literature, one approach is to transform the data to normality using the Johnson transform algorithm (Osanaiye *et al*., 2001) and thereafter, the $\bar{\mathrm{X}}$ and S control charts will be used to monitor the process stability of the transformed data instead of the original data. However, transformation of

data could lead to loss of information that may affect the inference drawn on such data (Shahriari *et al*., 2009). Abu-Shawiesh (2008) presented a simple approach to robust estimation of the process standard deviation based on the median absolute deviation. The Median Absolute Deviation (MAD) from the sample median is considered one of the good robust estimate of scale because it has high efficiency, high breakdown point, redescending influence function and how gross- error sensitivity which measures the worst influence a small amount of contamination of fixed size can have on the value of the estimator (Simonoff, 1984; Lax, 1985). Abu-Shawiesh (2008) considered the MAD as a robust estimator of the standard deviation when the sample data are not normal and then developed the MAD control chart for monitoring process variability. In this study, the concept of Abu-Shawiesh (2008) is extended to develop the control limits for the corresponding $\bar{\mathrm{X}}$ chart for monitoring the process average. Thus a combine control chart called $\bar{\mathrm{X}}$ and MAD chart is proposed in this study for monitoring variable quality characteristics when the normality assumption is not met.

Corresponding Author: Adekeye, K.S., Department of Mathematical Sciences, College of Natural Sciences,
Redeemers University (RUN), Nigeria

MATERIALS AND METHODS

$\overline{X}$ and Median Absolute Deviation (MAD) Charts

Let $x_{i1}, x_{i2},\ldots, x_{in}$ be a random sample of size n of independent observations taken at period i, where

i = 1, 2,…, m. The samples are assumed to be equal, independent and taken from continuous identical distribution functions. The robust control chart based on the MAD estimator is a chart of subgroup standard deviations (S_i) in which the control limits for the sake of robustness are set using the median absolute deviation from the sample median.

The Median Absolute Deviation from the sample median (MAD) is a very robust scale estimator than the sample standard deviation (Abu-Shawiesh, 2008). It measures the deviation of the data from the sample median. The MAD is often used as an initial value for the computation of more efficient robust estimators. Let x_1, x_2, x_n be a simple random sample of size n observation taken over m subgroups, then the MAD is defined as Eq. 1:

$$MAD = 1.4826MD\ \{|X_i - MD|\} \qquad (1)$$

where, MD is the median of x_1, x_2, …………., x_n. The average of the MAD is computed using Eq. 2:

$$\overline{MAD} = \sum_{i=1}^{m} MAD / m \qquad (2)$$

The control limits for the Shewhart S-control chart based on MAD as derived by Abu-Shawiesh (2008) is:

$$\begin{aligned} LCL &= B^*_5 \overline{MAD} \\ CL &= C^*_4 \overline{MAD} \\ UCL &= B^*_6 \overline{MAD} \end{aligned} \qquad (3)$$

where, $C_4 = \dfrac{\overline{S}}{\sigma} = \dfrac{4(n-1)}{4n-3}$, $B^*_5 = b_n(C_4 - 3\sqrt{1 - C^2_4})$

and $B^*_6 = b_n(C_4 + 3\sqrt{1 - C^2_4})$. It should be noted that the value of b_n, B^*_5 and B^*_6 have been derived and tabulated in Abu-Shawiesh (2008).

In order to take decision, the values of S_i, i = 1, 2,….., m, are plotted along with the values of control limits in Eq. 3 Materials and Method section is missing to obtain the MAD control chart. If any of the plotted S_i's falls outside the control limits in Eq. 3, the process variability is adjudged to be non stable.

Derivation of Control Limits For $\overline{X}$.

Let x_1, x_2, …, x_n be a simple random sample of size n observation taken over m subgroups. To derive the control limits for the corresponding $\overline{X}$ control chart for the MAD based control chart, we use the 3-sigma approach. The control limits are derived as follows:

Table 1: Number of Out-of-Control Points and Control Limits Interval (The values in the parentheses represent the control limits interval)

Distribution	$\overline{X}$ and S control charts		$\overline{X}$ and MAD control charts	
Standard Normal	0 (1.93)	0 (1.42)	0 (6.08)	0 (1.41)
Normal	0 (1.84)	0 (1.35)	0 (5.88)	0 (1.37)
Exponential	0 (3.50)	2 (2.57)	0 (7.34)	8 (1.70)
Weibul	0 (6.10)	7 (4.48)	0 (5.45)	22 (1.27)
Chi-Square	0 (2.29)	3 (1.68)	0 (4.67)	9 (1.08)

The 3- sigma control limits (Upper and Lower) are usually defined as Eq. 4:

$$\begin{aligned} UCL &= \overline{X} + 3\hat{\sigma}_x \\ CL &= \overline{X} \\ LCL &= \overline{X} - 3\hat{\sigma}_x \end{aligned} \qquad (4)$$

When MAD is used as an estimate of variability, then MAD will be used as a replacement of the standard deviation (S), Thus, $\hat{\sigma}_x = b_n \overline{MAD}$ (Abu-Shawiesh, 2008).

Then, the control limits in Eq. 4 will become:

$$\begin{aligned} UCL &= \overline{\overline{X}} + 3b_n \overline{MAD} \\ CL &= \overline{\overline{X}} \\ LCL &= \overline{\overline{X}} - 3b_n \overline{MAD} \end{aligned} \qquad (5)$$

Let $A_5 = 3b_n$, then the control limits in Eq. 5 will be reduced to Eq. 6:

$$\begin{aligned} UCL &= \overline{\overline{X}} + A_5 \overline{MAD} \\ CL &= \overline{\overline{X}} \\ LCL &= \overline{\overline{X}} - A_5 \overline{MAD} \end{aligned} \qquad (6)$$

Using the values of b_n as derived by Abu-Shawiesh (2008), the values of A_5 are computed and are presented in Table 1 in the appendix for various values of n. In this study an algorithm was developed and implemented on S-Plus to evaluate the performance of the proposed control chart with the traditional $\overline{X}$ and S charts.

RESULTS

Real life data: Two real life data from two different manufacturing companies in Nigeria are considered in this work for implementation to study the performance of the proposed $\overline{X}$ and MAD control charts. These are data on the hardness of Cofta tablets and the Brix content of Pepsi. The hardness of Cofta tablet obtained

on 30 days production with 10 batches of production per day as recorded in the quality control section of a manufacturing company in Nigeria is used as case1 while the Brix content of Pepsi obtained over 31 days production with 20 batches of production per day as recorded in the quality control section of a soft drink production company in Nigeria is used as case 2. Thus, the total number of observation for the hardness characteristic is 300 while the total number of observation for the Brix content is 620.

Case 1: Hardness of Cofta Tablet: In order to monitor the hardness of Cofta tablet, we first test the data of the hardness for the assumption of normality. Using the Shapiro and Wilk (1965) test for normality, the data on hardness yield the following:

Shapiro Wilk W = 0.895 and PW = 0.000. Therefore, the data is adjudged to be non-normal. Hence, the proposed $\bar{X}$ and MAD control charts will be an appropriate control chart to monitor the stability of the process. The values of the $\bar{X}$, S and MAD for the data were computed for each subgroup and then used to derived the control limits. From the data:

$$\overline{MAD} = \frac{86.1465}{30} = 2.87$$

and:

$$\bar{\bar{X}} = \frac{452.5}{30} = 15.08$$

From Table 2, $b_n = 1.087$ and $A_5 = 3.261$ when n = 10. Other control factors can be obtained from the SQC Table (Hampel, 1974; John *et al.*, 2006; Montgomery, 2008). Thus, the control limits for the $\bar{X}$ and MAD charts are obtained using Eq. 3 and 6. The obtained values of the $\overline{MAD}$, $\bar{\bar{X}}$, b_n and A_5 were substituted into Eq. 6 to derived the control limits for the $\bar{X}$ -chart as shown below:

$$UCL_{\bar{x}} = 15.08 + 3.261(2.87) = 24.44$$
$$CL_{\bar{x}} = 15.08$$
$$LCL_{\bar{x}} = 15.08 - 3.261(2.87) = 5.72$$

Similarly, the control limits for the MAD- control chart using Eq. 3 are:

$$UCL_{MAD} = 1.814 * 2.87 = 5.21$$
$$CL_{MAD} = 1.057 * 2.87 = 3.03$$
$$LCL_{MAD} = 0.300 * 2.87 = 0.86$$

Table 2: The Control limits factors for the proposed $\bar{X}$ and MAD control charts

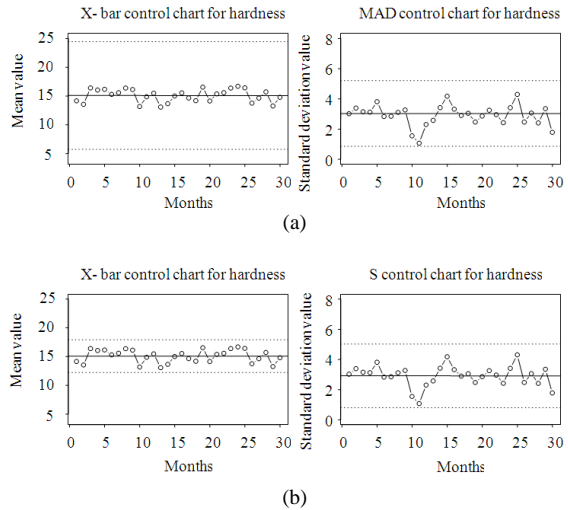| n	b_n	A_5
2	1.196	3.588
3	1.495	4.485
4	1.363	4.089
5	1.206	3.618
6	1.200	3.600
7	1.140	3.420
8	1.129	3.387
9	1.107	3.321
10	1.087	3.261
11	1.078	3.234
12	1.071	3.213
13	1.066	3.198
14	1.061	3.183
15	1.056	3.168
16	1.053	3.159
17	1.049	3.147
18	1.047	3.141
19	1.044	3.132
20	1.042	3.126
21	1.040	3.120
22	1.038	3.114
23	1.036	3.108
24	1.034	3.102
25	1.033	3.099

Fig. 1: $\bar{X}$ and MAD and $\bar{X}$ and S control charts for hardness

Using the computed control limits for the two control charts and the values of $\bar{X}$, S and MAD, the control charts for the hardness of Cofta tablet is presented in Fig. 1.

From Fig. 1, it is clear that both the process average and the process variability are in statistical control. Thus, the process that produces the hardness of Cofta tablet is said to be stable (i.e., the existing variation in the process average and variability are due

to random variation). It should be noted that the results on the $\overline{X}$ and *S* control chart do not differ significantly from the proposed chart (Fig. 1). However, there exist significant difference in the control limit intervals of the $\overline{X}$ control chart.

Case 2: Brix content of soft drink: To monitor the stability of the Brix content of Pepsi soft drink, we first test for the assumption of normality using the Shapiro and Wilk (1965) test for normality. The results obtained show that Shapiro and Wilk W = 0.769 and PW = 0.000. Therefore, the data is adjudged to be non-normal. Hence, the proposed $\overline{X}$ and MAD control charts will be an appropriate control chart to monitor the stability of the process. From the data, $\overline{MAD}$ = 0.1423 and $\overline{X}$ =10.8633. From Table 2, when n is equal to 20,

b_n = 1.042 and A_5 =3.126 . Also from the SQC Table, when n is equal to 20, B_5 = 0.504, B_6 =1.470 and A_3 =0.660. Using Eq. 6, the control limits for the $\overline{X}$ charts are UCL_x = 11.308 and LCL_x = 10.4185. Similarly, using Eq. 3, the control limits for the MAD-chart are

UCL_{MAD} = 0.218 and LCL_{MAD} = 0.075. The plot of the mean ($\overline{X}$) values and the standard deviations (S) values along with the computed control limits are presented in Fig. 2.

From Fig. 2, the following observations were made:

- The MAD based $\overline{X}$ control chart differ significantly from the S based $\overline{X}$ control chart. There exist no points on the MAD based $\overline{X}$ control chart that were outside the control limits while three (3) points were outside the limits on the S based $\overline{X}$ control chart
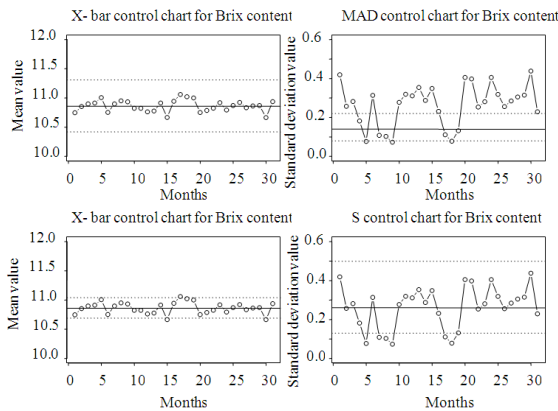

Fig. 2: $\overline{X}$ and MAD and $\overline{X}$ and S control charts for brix content of soft drink

- The control limits interval for the MAD based $\overline{X}$ control chart is higher than that of the S based $\overline{X}$ control charts
- On the MAD control chart about 87% of the plotted points were outside the control limits where as only 20% were outside the control limits on the S control chart

From the aforementioned points, it is clear that the use of a wrong control chart will surely lead to misplacement of control limits and hence high rate of false alarm.

Simulated data: To evaluate the performance of the $\overline{X}$ and MAD control chart, 300 data sets were generated and arranged in 10 sample size with 30 subgroups using the Normal, Exponential, Chi-square and Weibul distributions. The mean, standard deviation and MAD for the generated data in 10 sample size with 30 subgroups were computed and implemented to obtain the control limits for the $\overline{X}$ and MAD and $\overline{X}$ and S control charts. The summary of the out-of control points and the control limits interval for the control charts considered in this study are presented in Table 1.

DISCUSSION

The results from the two real life data (non-normal data) show that the $\overline{X}$ chart whether used with standard deviation (S) or Median Absolute Deviation (MAD) reflects that the process averages were in statistical control. However, there exists a small difference in the control limits interval. For instance, the control limits interval of the $\overline{X}$ control charts for the hardness of Cofta is 18.72 when the median absolute deviation was used to estimate the $\overline{X}$ control limits while it is 5.70 when the standard deviation was used. Thus, the control limits interval is higher when the median absolute deviation was used to compute the control limits of the $\overline{X}$ control chart than when the standard deviation was used. The import of the result is that, with the standard deviation used in computing the control limits, there is the possibility that some values which are within the control limits with the MAD based control limits will be outside the control limits (Fig. 1 and 2). Hence, misplacement of control limits is probable when the standard deviation is used to derive the $\overline{X}$ control limits for monitoring non-normal data. For monitoring the variability of the process, it was observed that the MAD based control limits detect changes in the variability of the process faster than the standard deviation based control limits. The results of the Brix content of soft drink and the generated data reflect similar results.

CONCLUSION

The results of the derived control limits for the $\bar{X}$ and MAD control chart are consistent with the existing methods. On comparing the $\bar{X}$ and S chart with the $\bar{X}$ and MAD chart, when used on a normal data, it was observed that both charts gave the same result. Thus, when the data under study is normal, either $\bar{X}$ and S chart or $\bar{X}$ and MAD chart can be used without any false alarm. However, when the $\bar{X}$ and MAD chart and the $\bar{X}$ and S chart were applied on a non-normal data, it was observed that only the MAD control charts shows that the process is not stable. This result is consistent with the Exploratory Data Analysis (EDA) results of the quality characteristic under study. The obtained results also reflect the possibility of misplacement of control limits if an appropriate control chart is not applied for monitoring purpose.

Furthermore, the derived control limits for the $\bar{X}$ control chart is an improvement on the control limits of the Shewhart $\bar{X}$ control chart since the control limits interval is higher than that of the standard deviation based control limits. Therefore, when the data under study is non-normal, the $\bar{X}$ and MAD control charts are more appropriate for monitoring the stability of the process. This is similar with the simulation result of Abu-Shawiesh (2008).

REFERENCES

Abu-Shawiesh, M.O.A., 2008. A simple robust control chart based on MAD. J. Math. Stat., 4: 102-107. DOI: 10.3844/JMSSP.2008.102.107

Adekeye, K.S., 2009. An Introduction to Industrial Statistics (Concepts and Practice). Adewumi.

Hampel, F.R., 1974. The Influence curve and its role in Robust estimation. J. Am. Stat. Assoc. 69: 383-393. DOI: 10.2307/2285666

John, J.A., D. Whitaker and D.G. Johnson, 2006. Statistical Thinking in Business. 2nd Edn., Chapman and Hall/CRC, Boca Raton, ISBN: 1584884959, pp: 394.

Lax, D.A., 1985. Robust estimators of scale: Finite-sample performance in long-tailed symmetric distributions. J. Am. Stat. Assoc., 80: 736-741. DOI: 10.2307/2288493

Montgomery, D.C., 2008. Introduction to Statistical Quality Control. 6th Edn., Wiley, Hoboken, N.J., ISBN: 0470169923, pp: 734.

Osanaiye, P.A., Adekeye, K.S. and Olugbara, O.O. 2001. Effect of normalizing statistical process control data using a developed johnson transformed-typed algorithm. J. Nigerian Stati. Assoc., 14: 34-51.

Shahriari, H., A. Maddahi and A.H. Shokouhi, 2009. A robust dispersion control chart based on M-estimate. J. Indus. Syst. Eng., 2: 297-307.

Shapiro, S.S. and M.B. Wilk, 1965. An analysis of variance test for normality (complete samples). Biometrika, 52: 591-611.

Simonoff, J.S., 1984. A comparison of robust methods and detection of outliers techniques when estimating a location parameter. Commun. Stat. Theory Meth., 13: 813-842. DOI: 10.1080/03610928408828722