

MODIFIED SIMPLE ROBUST CONTROL CHART BASED ON MEDIAN ABSOLUTE DEVIATION

Kayode Samuel Adekeye, PhD

Department of Mathematical Sciences, Redeemer's University,
Ogun State, Nigeria

Abstract:

The control limits derived for the Median Absolute Deviation (MAD) based Standard deviation (S) control chart proposed by Abu-Shawiesh was for monitoring quality characteristics when a standard value of sigma (σ) is known or given by the management/engineers. When sigma (σ) is unknown and we are interested in monitoring past/non-normal data, then there is the need to modify the simple robust control limits. In this paper, the control limits for the Shewhart $\bar{x}$ and S control chart based on median absolute deviation were modified using the concept of three sigma (3σ) limits. An evaluation performance tool was also developed to evaluate the efficiency of the modified control chart. An algorithm implemented on S-Plus programming language was developed to compute the two evaluation parameters used in this study. The results show that the control limits interval and the average run length for the modified control charts is smaller than that of the existing control charts. Therefore, the modified control limits is more efficient than the existing control limits. It is recommended that the modified control limits be used when monitoring past/non-normal data or when there is no standard value of sigma specify by the process engineer/ management.

Key Words: Median Absolute Deviation (MAD), Average Run Length (ARL), Control Limits Interval (CLI), Process control, Sigma approach

Introduction:

Statistical process control is commonly used as a tool in improving product quality through the achievement of process stability and capability. Monitoring and reducing variability in the process is a goal of statistical process control. With control charts, one of an on-line process control technique, changes of quality characteristics caused by common causes or assignable causes will be investigated. Then further corrective actions can be taken in order to remove those causes and return the process to the target with stable operation. Determination of the common or assignable causes of variation in control chart is possible

with the use of control limits. Therefore, incorrect estimation of the control limits will lead to incorrect inference and thus wrong action.

Specifying the control limits is the most important step in designing a control chart. Improper estimation of the process dispersion which results in narrower or wider limits can increase the probability of type I error or the probability of type II error. When the limits are narrow, the risk of a point falling beyond the limits increases, falsely indicating that the process is out of control (Shahriari et al., 2009). When the limits are wider the risk increases the points falling within the limits, this falsely indicates that the process is out of control (Shahriari et al., 2009). The limits on a control chart can either be 0.001 probability limits or 3-sigma limits (Oakland 2008). The 0.001 probability limit were determined so that, if chance causes alone were at work the probability of a point falling above the upper limit would be one out of a thousand and the probability of points falling below the lower limit would be one out of a thousand. Therefore, 3- sigma limits are the practical equivalent of the 0.001 probability limits. In this paper, the 3-sigma limits approach will be adopted.

Abu-Shawiesh (2008) derived the control limits for the standard deviation control chart using the median absolute deviation. However, the derived control limits was for situation when a standard value of sigma is known or specified by the management. Using such control limits to monitor a process where the standard value of sigma is unknown will lead to misplacement of control limits and hence false alarm of out of control. In this paper, we consider situation when the standard value of sigma is unknown and we are interested in monitoring past data. Thus, we modify the control limits earlier presented by Abu-Shawiesh (2008) for the S-control chart and Adekeye et al. (2012) for the $\bar{X}$ control chart

2.0 Materials and Methods

Let X_{ij} represent a random sample of size n taken over m subgroup, $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$. The sample are assumed to be independent and taken from a continuous identical distribution functions. If σ^2 is unknown, then an unbiased estimate of σ^2 is the sample variance (S^2) evaluated by

$$S_j^2 = \frac{1}{n-1} \left[\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \right],$$

Where $\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{ij}$.

If the underline distribution is normal, then an unbiased estimator of S is $C_4\sigma$, where $C_4 = (2/n-1)^{1/2} [\Gamma(n/2)/\Gamma((n-1)/2)]$ and the standard deviation of S is $\sigma\sqrt{1-C_4^2}$. That is, $E(S) = C_4 \sigma$ and $V(S) = \sigma^2(1-C_4^2)$. Therefore, the 3-sigma control limits for S- chart will be

$$\begin{aligned} UCL &= C_4\sigma + 3\sigma\sqrt{1-C_4^2} \\ CL &= C_4\sigma \\ LCL &= C_4\sigma - 3\sigma\sqrt{1-C_4^2} \end{aligned} \quad (1)$$

For simplification of Equation (1) , let $B_5 = C_4 - 3\sqrt{1-C_4^2}$ and $B_6 = C_4 + 3\sqrt{1-C_4^2}$..

Then the parameter of Equation (1) becomes

$$\begin{aligned} UCL &= B_6\sigma \\ CL &= C_4\sigma \\ LCL &= B_5\sigma \end{aligned} \quad (2)$$

Suppose we use $\hat{\sigma} = \frac{\bar{S}}{C_4}$ which is an unbiased estimator of σ , then the control limits for the S- chart in Equation (1) will be reduced to

$$\begin{aligned} UCL &= \bar{S} + 3(\bar{S}/C_4)\sqrt{1-C_4^2} \\ CL &= \bar{S} \\ LCL &= \bar{S} - 3(\bar{S}/C_4)\sqrt{1-C_4^2} \end{aligned} \quad (3)$$

Where $\bar{S} = \frac{1}{m} \sum_{j=1}^m S_j$. For simplification of Equation (3), let $B_3 = 1 - (3/C_4)\sqrt{1-C_4^2}$ and

$B_4 = 1 + (3/C_4)\sqrt{1-C_4^2}$, then the control limits for S- Chart for analyzing past data in

Equation (3) becomes

$$\begin{aligned} UCL &= B_4\bar{S} \\ CL &= \bar{S} \\ LCL &= B_3\bar{S} \end{aligned} \quad (4)$$

To monitor the process average, the 3-sigma control limits for monitoring past data (retrospective analysis) are

$$\begin{aligned}
 UCL &= \bar{\bar{X}} + 3 \frac{\sigma}{\sqrt{n}} \\
 CL &= \bar{\bar{X}} \\
 LCL &= \bar{\bar{X}} - 3 \frac{\sigma}{\sqrt{n}}
 \end{aligned}
 \tag{5}$$

Where $\bar{\bar{X}} = \frac{1}{m} \sum_{j=1}^m \bar{X}_j$. When $\bar{S}/C_4$ is used to estimate σ , then the control limits on the corresponding $\bar{X}$ - charts will be

$$\begin{aligned}
 UCL &= \bar{\bar{X}} + 3 \frac{\bar{S}}{C_4 \sqrt{n}} \\
 CL &= \bar{\bar{X}} \\
 LCL &= \bar{\bar{X}} - 3 \frac{\bar{S}}{C_4 \sqrt{n}}
 \end{aligned}
 \tag{6}$$

To simplify Equation (6), let $A_3 = 3 / C_4 \sqrt{n}$, then Equation (6) is reduced to:

$$\begin{aligned}
 UCL &= \bar{\bar{X}} + A_3 \bar{S} \\
 CL &= \bar{\bar{X}} \\
 LCL &= \bar{\bar{X}} - A_3 \bar{S}
 \end{aligned}
 \tag{7}$$

The constants B_3 , B_4 , and A_3 are obtainable from the appropriate SQC Table for various value of n . The above control limits are for situation when the process to be monitored is normal. In practice, the normality assumption is often violated by real life data, therefore, using the above control limits to monitor real life data might leads to wrong signal and invalid inference. The median absolute deviation (MAD) has been claimed in the literature to be the best estimate when the data under consideration is non-normal. When MAD is used as an estimate of variability, then

$$\hat{\sigma} = b_n \overline{MAD} \tag{8}$$

Where $\overline{MAD} = \frac{\sum_{j=1}^m MAD_j}{m}$, b_n is a function of the sample size n and

$$MAD_j = 1.4826 [\text{Median} \{ |X_{ij} - MD_j| \}] \tag{9}$$

Where $MD_j = \text{Median} (X_{ij})$

Thus, the control limits and the central line for the Shewhart S- control chart based on the Median Absolute Deviation as proposed by Abu-Shawiesh (2008) were obtained by

substituting Equation (8) into Equation (2). Therefore, the S- control chart limits based on MAD are

$$\begin{aligned} UCL_s &= B_6 b_n \overline{MAD} \\ CL_s &= C_4 b_n \overline{MAD} \\ LCL_s &= B_5 b_n \overline{MAD} \end{aligned} \quad (10)$$

where, $B_5 = C_4 - 3\sqrt{1 - C_4^2}$ and $B_6 = C_4 + 3\sqrt{1 - C_4^2}$

However, if we are interested in monitoring past data, i.e. carrying out a retrospective analysis on a process data, the above expression in equation (10) will lead to misplacement of control limits. To solve this problem, the expression in Equation (4) will be used. On substituting the MAD based estimate of sigma in Equation (8) into Equation (4), then we have

$$\begin{aligned} UCL_s &= B_4 b_n \overline{MAD} \\ CL &= b_n \overline{MAD} \\ LCL_s &= B_3 b_n \overline{MAD} \end{aligned} \quad (11)$$

The corresponding control limits for the $\bar{X}$ charts depending on whether the sigma is known or unknown is consider as follow.

If σ is known, then the derived control limits by Adekeye (2011) will be used. However, if σ is unknown, then the control limits will be derived from Equation (5). Substituting Equation (8) into Equation (5), we have

$$\begin{aligned} UCL &= \bar{\bar{X}} + 3 \frac{b_n \overline{MAD}}{\sqrt{n}} \\ CL &= \bar{\bar{X}} \\ LCL &= \bar{\bar{X}} - 3 \frac{b_n \overline{MAD}}{\sqrt{n}} \end{aligned} \quad (12)$$

Let $A_6 = \frac{3 \bullet b_n}{\sqrt{n}}$, then the control limits in Equation (12) will be reduced to:

$$\begin{aligned} UCL &= \bar{\bar{X}} + A_6 \overline{MAD} \\ CL &= \bar{\bar{X}} \\ LCL &= \bar{\bar{X}} - A_6 \overline{MAD} \end{aligned} \quad (13)$$

In general, if the data under consideration is non-normal and there is no management specification, then the control limits in Equations (13) and (11) should be used to monitor the

process mean and variability, respectively. Table 1 gives value of the constants (b_n , A_5 and A_6) needed for the computation of the modified control limits in Equations (11) and (13). It should be noted that the value of the constants B_3 , B_4 , and C_4 can be obtained in any standard statistical quality control text book.

Table 1: Derived Constants For Computation of The Proposed Control Limits

n	b_n	A_5	A_6
2	1.196	3.588	2.5371
3	1.495	4.485	2.5893
4	1.363	4.089	2.0445
5	1.206	3.618	1.6180
6	1.200	3.600	1.4697
7	1.140	3.420	1.2926
8	1.129	3.387	1.1975
9	1.107	3.321	1.1070
10	1.087	3.261	1.0312
11	1.078	3.234	0.9751
12	1.071	3.213	0.9275
13	1.066	3.198	0.8870
14	1.061	3.183	0.8507
15	1.056	3.168	0.8180
16	1.053	3.159	0.7897
17	1.049	3.147	0.7633
18	1.047	3.141	0.7403
19	1.044	3.132	0.7185
20	1.042	3.126	0.6990
21	1.040	3.120	0.6808
22	1.038	3.114	0.6639
23	1.036	3.108	0.6481
24	1.034	3.102	0.6332
25	1.033	3.099	0.6198

2.1 Evaluation Parameters

The most universal and applicable way to measure performance of control charts is the use of Average Run Length (ARL). In this study, the evaluation parameters used to evaluate the control charts performance is the control limit interval and average run length.

The control limit interval hereafter refer to as CLI, is the difference between the control limits value. Therefore, for the S-control charts, the control limit interval will be determined using the expression:

$$CLI_S = \left(\frac{6[\sqrt{1-C_4^2}]}{C_4} \right) b_n \overline{MAD} \quad (14)$$

Similarly, the control limit interval for the $\bar{x}$ control chart will be determined using the expression:

$$CLI_{\bar{x}} = \left[\frac{6b_n}{\sqrt{n}} \right] \overline{MAD} \quad (15)$$

A control chart will be adjudged to be superior if its Control Limit Interval is the smallest out of all the control charts under consideration.

3.0 Results

To evaluate the performance of the $\bar{x}$ 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 using the derived control limits in Equation (11) for S-control chart and Equation (13) for the $\bar{x}$ control charts. Using Equation (14) and Equation (15), the control limits interval was computed for the control charts under consideration. The summary of the Control Limits Interval (CLI) and the Average Run Length (ARL) for the $\bar{x}$ and S control charts using the derived control limits and the control limits of Abu-Shawiesh (2008) and Adekeye et al. 2012 are presented in Table 2

Table 2: Average Run Length (ARL) and Control Limits Interval (CLI)

Distribution of Data Set	$\bar{X}$ and S Control Limits				Modified $\bar{X}$ and S Control Limits			
	$\bar{X}$		S		$\bar{X}$		S	
	ARL	CLI	ARL	CLI	ARL	CLI	ARL	CLI
Standard Normal	419.80	6.08	1.53	1.41	2.96	1.92	1.53	1.45
Normal	305.62	5.88	1.00	1.37	2.84	1.86	1.52	1.40
Exponential	284.05	7.34	2.40	1.71	1.66	2.32	1.40	1.75
Weibull	8.80	5.95	1.46	1.38	1.27	1.88	1.46	1.42
Chi-Square	14.71	4.85	1.64	1.13	2.06	1.53	1.65	1.16

4.0 Discussion of Results

The evaluation parameters used in this study (the control limits interval and the average run length) show insignificant difference for the S control charts when compare with the earlier results by Abu-shawiesh (2008) and Adekeye *et al.* (2011). However, for the $\bar{x}$ control charts, there is considerable improvement as seen in the control limits interval and the average run length for all the data set considered in this study. Furthermore, it can be seen that the average run length decreases as the data set deviates from normality. This is an indication that the median absolute deviation is more efficient and hence more appropriate than the standard deviation for highly skewed data. Thus, researcher or control chart users are advised to first carry out an exploratory data analysis on process data (real life data) before the use of control chart. This will help to guide against misplacement of control limits.

5.0 Conclusion

We have modified the simple robust control limits for monitoring process data when the standard value of sigma is not given and Median Absolute Deviation (MAD) is used to give an estimate. An S-plus programme has also been presented for the proposed control limits and the computation of control limit interval and average run length (see the Appendix). From the results in Table 2, it is clear that the average run length for the $\bar{x}$ chart for the modified control limits is more efficient than the earlier control limits for all the distributional data sets under consideration. Therefore, the modified control limits derived in this work for the $\bar{x}$ chart can be used when monitoring past/non-normal data or when there is no standard value of sigma specify by the process engineer/ management.

References:

- Abu-Shawiesh, M.O.A. (2008). A Simple Robust Control Chart Based on MAD. *Journal of Mathematics and Statistics* 4(2): 102 – 107. DOI: 10.3844/JMSSP.2008.102.107
- Adekeye K. S. and Azubuike, P.I. (2012). Derivation of the Limits For Control Chart Using The Median Absolute Deviation For Monitoring Non –Normal Process. *Journal of Mathematics and Statistics* 8(1): 37 – 41. DOI: 10.3844/JMSSP.2012.37.41
- Oakland, J. S., (2008). *Statistical Process Control*, Sixth Edition, Butterworth- Heinemann, London.
- Shahriari H., Maddahi A., Shokouhi A.H. (2009). A Robust Control Chart Based on M-Estimate. *Journal of Industrial and Systems Engineering* 2(4):. 297 – 307.

Appendix: S-Plus Program for Evaluation Parameters and Control Limits Constant

```
msim.data<-function(data){  
  A3<-0.975  
  B4<-1.716  
  B3<-0.284  
  A5.10<-3.261  
  bn.10<-1.087  
  C4<-0.9727  
  A6.10<-1.0312  
  POC<-NULL  
  mean.data<-apply(data,1, mean)  
  sta.data<-sqrt(apply(data,1,var))  
  mad.data<-sqrt(apply(data,1,mad))  
  avl.mean<-mean(mean.data)  
  avl.med<-mean(mad.data)  
  avl.sta<-mean(sta.data)  
  UCL.XbarS<-avl.mean + (A3*avl.sta)  
  CL.XbarS<-avl.mean  
  LCL.XbarS<-avl.mean- (A3*avl.sta)  
  UCL.Staad<-B4*avl.sta  
  CL.Staad<-avl.sta  
  LCL.Staad<-B3*avl.sta  
  diffXS<-UCL.XbarS-LCL.XbarS  
  diffS<-UCL.Staad-LCL.Staad  
  MUCL.XbarM<-avl.mean+(A6.10*avl.med)  
  MCL.XbarM<-avl.mean  
  MLCL.XbarM<-avl.mean-(A6.10*avl.med)  
  MUCL.mad<-B4*bn.10*avl.med  
  MCL.mad<-1.057*avl.med  
  MLCL.mad<-B3*bn.10*avl.med
```

```

CLI.Xm<-(MUCL.XbarM-MLCL.XbarM)
CLI.Sm<-(MUCL.mad-MLCL.mad)
if (mean.data[mean.data >UCL.XbarS] | mean.data[mean.data <LCL.XbarS]){
  POC<-c(POC, 1)}
else{
  POC<-c(POC, 0)}
MPOC.Xbars <- c(mean.data[mean.data >UCL.XbarS],mean.data[mean.data <LCL.XbarS])
MPOC.staad <- c(sta.data[sta.data>UCL.Staad],sta.data[sta.data<LCL.Staad])
MPOC.xbarM <- c(mean.data[mean.data >MUCL.XbarM],mean.data[mean.data
  <MLCL.XbarM])
MPOC.mad <- c(sta.data[sta.data >MUCL.mad],sta.data[sta.data <MLCL.mad])
RL.sn<-1-((pnorm(mdata1.gen$MUCL.XbarM,0,1))-pnorm(mdata1.gen$MLCL.XbarM,0,1))
ARL.sn<-1/RL.sn
RL.snm<-1-((pnorm(mdata1.gen$MUCL.mad,0,1))-pnorm(mdata1.gen$MLCL.mad,0,1))
ARL.snm<- 1/RL.snm
ARL.snm
return(MPOC.xbarM,MPOC.mad,CLI.Xm,CLI.Sm)

```