



Science

SAMPLE STANDARD DEVIATION(s) CHART UNDER THE ASSUMPTION OF MODERATENESS AND ITS PERFORMANCE ANALYSIS

Kalpesh S. Tailor ^{*1}

^{*1} Assistant Professor, Department of Statistics, M. K. Bhavnagar University, Bhavnagar-364001, India

DOI: <https://doi.org/10.5281/zenodo.821537>

Abstract

Moderate distribution proposed by Naik V.D and Desai J.M., is a sound alternative of normal distribution, which has mean and mean deviation as pivotal parameters and which has properties similar to normal distribution. Mean deviation (δ) is a very good alternative of standard deviation (σ) as mean deviation is considered to be the most intuitively and rationally defined measure of dispersion. This fact can be very useful in the field of quality control to construct the control limits of the control charts. On the basis of this fact Naik V.D. and Tailor K.S. have proposed 3δ control limits. In 3δ control limits, the upper and lower control limits are set at 3δ distance from the central line where δ is the mean deviation of sampling distribution of the statistic being used for constructing the control chart. In this paper assuming that the underlying distribution of the variable of interest follows moderate distribution proposed by Naik V.D and Desai J.M, 3δ control limits of sample standard deviation(s) chart are derived. Also the performance analysis of the control chart is carried out with the help of OC curve analysis and ARL curve analysis.

Keywords: Mean Deviation; Standard Deviation; Control Charts; Normal Distribution; Moderate Distribution; Statistic; OC Function; ARL Curve.

Cite This Article: Kalpesh S. Tailor. (2017). "SAMPLE STANDARD DEVIATION(s) CHART UNDER THE ASSUMPTION OF MODERATENESS AND ITS PERFORMANCE ANALYSIS." *International Journal of Research - Granthaalayah*, 5(6), 368-377. <https://doi.org/10.5281/zenodo.821537>.

1. Introduction

A fundamental assumption in the development of control charts for variables is that the underlying distribution of the quality characteristic is normal. The normal distribution is one of the most important distributions in the statistical inference in which mean (μ) and standard deviation (σ) are the parameters. Naik V.D and Desai J.M. have suggested an alternative of normal distribution, which is called moderate distribution. In moderate distribution mean (μ) and

mean deviation (δ) are the pivotal parameter and, which has properties similar to normal distribution.

Naik V.D. and Tailor K.S. have proposed the concept of 3δ control limits on the basis of moderate distribution. Under this rule, the upper and lower control limits are set at 3δ distance from the central line where δ is the mean deviation of sampling distribution of the statistic being used for constructing the control chart. Thus in the proposed control charts, under the moderateness assumption, three control limits for any statistic T should be determined as follows.

$$\begin{aligned} \text{Central line (CL)} &= \text{Expected value of T} = \mu \\ \text{Lower Control Limit (LCL)} &= \text{Mean of T} - 3\delta_T = \mu - 3\delta_T \\ \text{Upper Control Limit (UCL)} &= \text{Mean of T} + 3\delta_T = \mu + 3\delta_T \end{aligned}$$

Where μ is mean of T and δ_T is the mean error of T.

It is found that since δ provides exact average distance from mean and σ provides only an approximate average distance, 3δ limits are more rational as compared to 3σ limits.

Therefore, in this paper it has been assumed that the underlying distribution of the variable of interest follows moderate distribution and sample standard deviation(s) chart is studied and its 3δ control limits are derived. Also the performance analysis of s chart under the assumption of moderateness against that of normality assumption is carried out through OC curve analysis and ARL curve analysis.

2. (3δ) Control Limits for Sample Standard Deviation(S) Chart When Process Mean Deviation δ' is Unknown

When the size of the sample is relatively large or it is required to draw a large sample from the production process or when it is required to control the process variation, the s-chart is used instead of R-chart. When the size of the sample is 10 or 15 or more, then the statistic R is not considered as an efficient estimator to measure the variation of the process. In such a case, sample standard deviation (s) is considered to be the best estimator. The objective and application of s-chart is same as the R-chart.

Suppose that the main variable of the process x follows **moderate distribution**. The mean of x is $E(x) = \mu$ and mean deviation of x is $\delta_x = \delta'$. For s chart the values of sample standard deviation(s) are obtained from each subgroup taken at regular interval of time, from a production process. As the value of process S.D. σ' is unknown, its estimator \bar{s} is used in its place.

$$\text{i.e. } E(s) = C_2\sigma' = \bar{s} \text{ or } \sigma' = \frac{E(s)}{C_2} = \frac{\bar{s}}{C_2}.$$

The following important results of σ (when the underlying distribution is moderate) are used for calculating the control limits of this chart.

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \text{ is the sampling variance} \quad (1)$$

$$\bar{s} = \frac{1}{m} \sum_{i=1}^m s_i \quad (2)$$

$$E(s) = \bar{s} \quad (3)$$

$$E\left(\frac{s}{\sigma'}\right) = C_2 \quad (4)$$

Where σ' is the process standard deviation and C_2 is a constant which depends on the size of the sample.

$$\therefore E(s) = C_2 \sigma' = \sqrt{\frac{\pi}{2}} C_2 \delta' \quad (5)$$

$$\text{Or } \sigma' = \frac{E(s)}{C_2} = \frac{\bar{s}}{C_2} \text{ and } \delta' = \frac{\bar{s}}{\sqrt{\frac{\pi}{2}} C_2} \quad (6)$$

Where δ' is the process mean deviation.

$$S.D\left(\frac{s}{\sigma'}\right) = [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{1}{\sqrt{2n}} \quad (7)$$

$$\therefore S.D(s) = [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{\sigma'}{\sqrt{2n}} \quad (8)$$

$$M.D(s) = \delta_s = [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \sqrt{\frac{\pi}{2}} \frac{\delta'}{\sqrt{2n}} \quad (9)$$

Hence on the basis of 3δ criteria the control limits of s chart can be represented as follows.

$$\begin{aligned} \text{Central line (C.L)} &= E(s) \\ &= \bar{s} \end{aligned} \quad (10)$$

$$\begin{aligned} \text{Lower control limit(L.C.L)} &= E(s) - 3\delta_s \\ &= \bar{s} - 3 \sqrt{\frac{2}{\pi}} \sigma_s \\ &= \bar{s} - 3 \sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{\sigma'}{\sqrt{2n}} \\ &= \bar{s} - 3 \sqrt{\frac{2}{\pi}} \left[[2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{\bar{s}}{C_2 \sqrt{2n}} \right] \\ \text{As } \sigma' &= \frac{E(s)}{C_2} = \frac{\bar{s}}{C_2} \\ &= \bar{s} \left\{ 1 - 3 \sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{1}{C_2 \sqrt{2n}} \right\} \\ &= B'_3 \bar{s} \end{aligned} \quad (11)$$

$$\text{Where } B'_3 = 1 - 3 \sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{1}{C_2 \sqrt{2n}}$$

$$\begin{aligned} \text{Upper control limit(U.C.L)} &= E(s) + 3\delta_s \\ &= \bar{s} + 3 \sqrt{\frac{2}{\pi}} \sigma_s \\ &= \bar{s} + 3 \sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{\sigma'}{\sqrt{2n}} \\ &= \bar{s} + 3 \sqrt{\frac{2}{\pi}} \left[[2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{\bar{s}}{C_2 \sqrt{2n}} \right] \\ \text{As } \sigma' &= \frac{E(s)}{C_2} = \frac{\bar{s}}{C_2} \\ &= \bar{s} \left\{ 1 + 3 \sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{1}{C_2 \sqrt{2n}} \right\} \end{aligned}$$

$$= B'_4 \bar{s} \tag{12}$$

$$\text{Where } B'_4 = 1 + 3 \sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{1}{C_2\sqrt{2n}}$$

Values of B'_3 , B'_4 for different values of n are given in the table 1.

Table 1:

Values of B'_3 and B'_4 for different values of n		
n	B'_3	B'_4
2	0	2.8080
3	0	2.2569
4	0	2.0099
5	0.1308	1.8691
6	0.2263	1.7737
7	0.2961	1.7039
8	0.3499	1.6501
9	0.3911	1.6088
10	0.4281	1.5719
11	0.4591	1.5409
12	0.4838	1.5162
13	0.5069	1.4931
14	0.5265	1.4735
15	0.5438	1.4562
16	0.5603	1.4397
17	0.5736	1.4264
18	0.5847	1.4152
19	0.5983	1.4017
20	0.6086	1.3914
21	0.6196	1.3804
22	0.6294	1.3706
23	0.6377	1.3623
24	0.6459	1.3541
25	0.6520	1.3480
13	0.5069	1.4931
14	0.5265	1.4735
15	0.5438	1.4562
16	0.5603	1.4397
17	0.5736	1.4264
18	0.5847	1.4152
19	0.5983	1.4017
20	0.6086	1.3914
21	0.6196	1.3804
22	0.6294	1.3706
23	0.6377	1.3623
24	0.6459	1.3541
25	0.6520	1.3480

3. Comparison of Performance of s Chart under the Assumption of Normality with 3σ- Limits Against that under the Assumption of Moderateness and 3δ-Limits

There are two commonly used methods for measuring and comparing the performance of control charts. One of them is to determine the Operating Characteristic (OC) curve of the charts and the other one is to determine the average run length (ARL).

It is very helpful to use the operating characteristic (OC) curve of a control chart to display its probability of type-II error. This would be an indication of the ability of the control chart to detect process shifts of different magnitudes. The OC Curve shows the probability that an observation will fall within the control limits given the state of the process. This is very much like finding power curves in hypothesis testing. Another measure of performance that is closely related to OC curve values is the run length. The run length is a random variable and is defined as the number of points plotted on the chart until an out-of-control condition is signaled.

4. Comparison of Performance of s Chart under the Assumption of Normality with 3σ- Limits against that under the Assumption of Moderateness with 3δ-Limits through Oc Curve Analysis

The two types of errors for control charts are defined as follows.

Let α = probability of type-I error of control charts.

= probability that the process is considered to be of out control when it is really in control, and

β = probability of type-II error of control charts.

= probability that the process is considered to be in control when it is really out of control, Clearly, $1 - \beta$ = probability of not committing type-II error on control charts. Thus lower value of β and higher value of $(1 - \beta)$ means more effectiveness (better performance) of control charts.

Consider the OC curve for an s chart with the known standard deviation σ . Suppose that the in-control value of standard deviation is σ_0 . If the standard deviation shifts from the in control value σ_0 to another value $\sigma_1 > \sigma_0$, then the probability of not detecting a shift to a new value of σ , i.e $\sigma_1 > \sigma_0$, by the first sample following the shift is,

$$\beta = P\{LCL \leq s \leq UCL / \sigma_1 > \sigma_0\} \quad (13)$$

Since mean of s is $\bar{s} = C_2 \sigma_0$ and mean error is $\sqrt{\frac{2}{\pi}} \sigma_1$ under the assumption of moderateness and since the upper and lower 3δ control limits are $UCL = B_4' \bar{s}$ and $LCL = B_3' \bar{s}$, the equation (13) can be written as follows,

$$\begin{aligned} \beta_{ms} &= \Phi' \left[\frac{UCL - \bar{s}}{\sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{\sigma_0}{\sqrt{2n}}} \right] - \Phi' \left[\frac{LCL - \bar{s}}{\sqrt{\frac{2}{\pi}} [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{\sigma_0}{\sqrt{2n}}} \right] \\ &= \Phi' \left[\frac{B_4' \bar{s} - \bar{s}}{C'_0 \sigma_1} \right] - \Phi' \left[\frac{B_3' \bar{s} - \bar{s}}{C'_0 \sigma_1} \right], \text{ where } C'_0 = [2(n-1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{1}{\sqrt{\pi n}} \\ &= \Phi' \left[\frac{\bar{s}(B_4' - 1)}{C'_0 \sigma_1} \right] - \Phi' \left[\frac{\bar{s}(B_3' - 1)}{C'_0 \sigma_1} \right] \\ &= \Phi' \left[\frac{C_2 \sigma_0 (B_4' - 1)}{C'_0 \sigma_1} \right] - \Phi' \left[\frac{C_2 \sigma_0 (B_3' - 1)}{C'_0 \sigma_1} \right] \end{aligned}$$

$$= \Phi' \left[\frac{C_2}{C'_0} \cdot \frac{(B_4' - 1)}{\lambda_2'} \right] - \Phi' \left[\frac{C_2}{C'_0} \cdot \frac{(B_3' - 1)}{\lambda_2'} \right] \quad (14)$$

Where $\lambda_2' = \frac{\sigma_1}{\sigma_0}$ or $\sigma_1 = \lambda_2' \sigma_0$ and Φ' denotes the standard moderate cumulative distribution.

Similarly, equation (22) can be written as follows under normality assumption,

$$\beta_{ns} = \Phi \left[\frac{C_2}{C_0} \cdot \frac{(D_4 - 1)}{\lambda_2} \right] - \Phi \left[\frac{C_2}{C_0} \cdot \frac{(D_3 - 1)}{\lambda_2} \right] \quad (15)$$

Where $C_0 = [2(n - 1) - 2nC_2^2]^{\frac{1}{2}} \cdot \frac{1}{\sqrt{2n}}$ and $\lambda_2 = \frac{\sigma_1}{\sigma_0}$

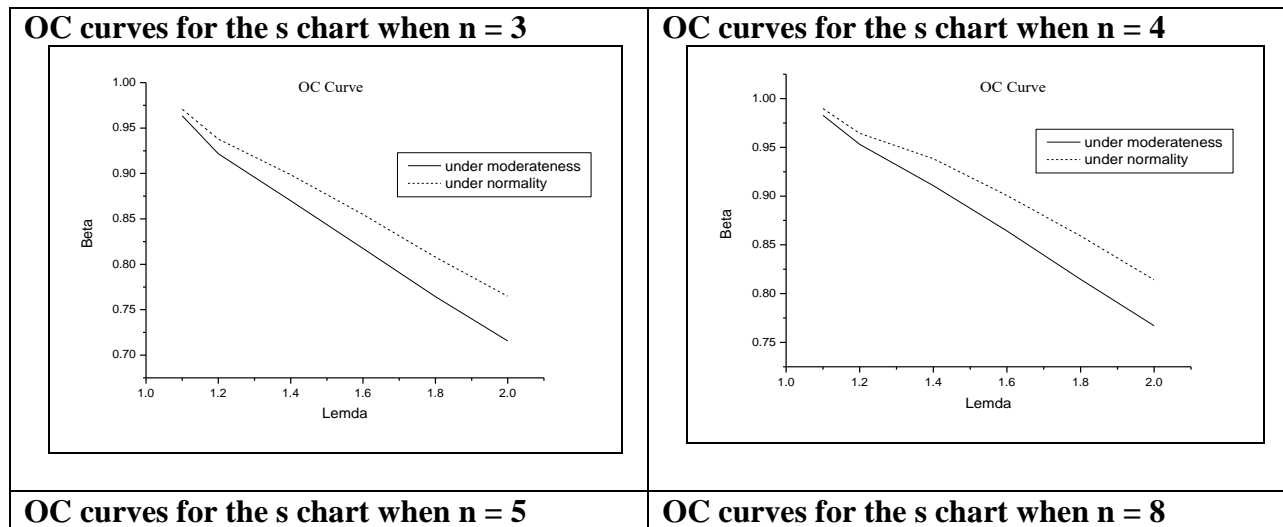
Usually shift in the value of δ_0 (or σ_0) is measured in terms of percentage of its 'in control' value. Thus $\delta_1 = 1.1\delta_0$ means 10% shift in the value of δ_0 , $\delta_1 = 1.5\delta_0$ means 50% shift in the value of δ_0 , $\delta = 2\delta_0$ means 100% shift in the value of δ_0 . Hence λ_1' (and λ_1) are usually chosen in the range [1, 2].

To construct the OC curve for s chart under moderateness (or normality) assumption, β -value is plotted against λ_2' (or λ_2) with various sample sizes n. These probabilities may be evaluated directly from equation (14) and (15).

For different sample sizes n and with three-delta limits (or three-sigma limits), for various values of λ_2' (or λ_2), β -values are calculated and OC curves are plotted as shown in figure-1.

Table 2:

λ_2 (or λ_2')	n = 3		n = 4		n = 5		n = 8	
	β_{ms}	β_{ns}	β_{ms}	β_{ns}	β_{ms}	β_{ns}	β_{ms}	β_{ns}
1.0	0.9634	0.9706	0.9828	0.9898	0.9834	0.9957	0.9834	0.9978
1.2	0.9217	0.9379	0.9531	0.9644	0.9540	0.9828	0.9540	0.9892
1.4	0.8699	0.8985	0.9107	0.9383	0.9122	0.9594	0.9122	0.9714
1.6	0.8175	0.8549	0.8643	0.9006	0.8664	0.9273	0.8664	0.9438
1.8	0.7643	0.8079	0.8146	0.8591	0.8172	0.8895	0.8172	0.9080
2.0	0.7157	0.7647	0.7671	0.8142	0.7686	0.8494	0.7686	0.8740



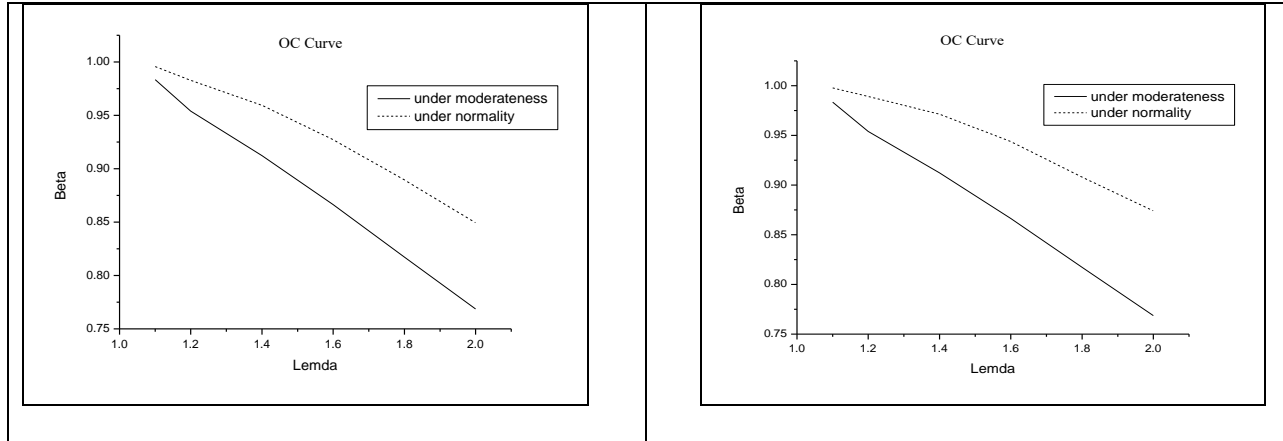


Figure 1:

Form the above table 3 and figure 1, it is noticeable that for all values of n when $\lambda_2' = \lambda_0'$ ($\lambda_2 = \lambda_0$), i.e even when there is no shift in the value of δ_0 (or σ_0), β under moderateness assumption are always smaller than β under normality assumption, which indicates that s-chart under moderateness assumption and 3δ -limits is more effective than that of normality assumption and 3σ -limits

5. Comparison of Performance of s Chart under the Assumption of Normality with 3σ - Limits against that of the Assumption of Moderateness with 3δ -Limits through ARL Curve Analysis

For a control chart the average run length (ARL) is the average number of points required to be plotted before a point indicates an out of control condition, that means when ARL is small, the chart is considered to be more effective. If the process observations are uncorrelated, then for any control chart, the ARL can be calculated easily from,

$$ARL = \frac{1}{p} = \frac{1}{1-\beta} \quad (16)$$

Where p is the probability that any point exceeds the control limits. This equation can be used to evaluate the performance of the control chart.

To construct the ARL curve for the s chart, ARL is plotted against the magnitude of the shift with various sample sizes n . To measure the effectiveness of the control charts under both the assumption, viz moderateness and normality, the probabilities (β) may be evaluated directly from equations (14) and (15) and values of ARL are calculated from equation (16) and ARL curves are plotted. For different sample sizes n and with 3δ limits (or 3σ limits), for various values of λ_2' (or λ_2), ARLs are calculated and ARL curves are plotted as shown below.

Table 3:

λ_2 (or λ_2')	n = 3		n = 4		n = 5		n = 8	
	ARL _{ms}	ARL _{ns}	ARL _{ms}	ARL _{ns}	ARL _{ms}	ARL _{ns}	ARL _{ms}	ARL _{ns}
1.0	27	34	58	98	60	233	60	455
1.2	13	16	21	28	22	58	22	93
1.4	8	10	11	16	11	25	11	35
1.6	5	7	7	10	7	14	8	18

1.8	4	5	5	7	5	9	5	11
2.0	4	4	4	5	4	7	4	8

Where ARL_{ms} = ARL values under moderateness assumption for s-chart.
 ARL_{ns} = ARL values under normality assumption for s-chart.

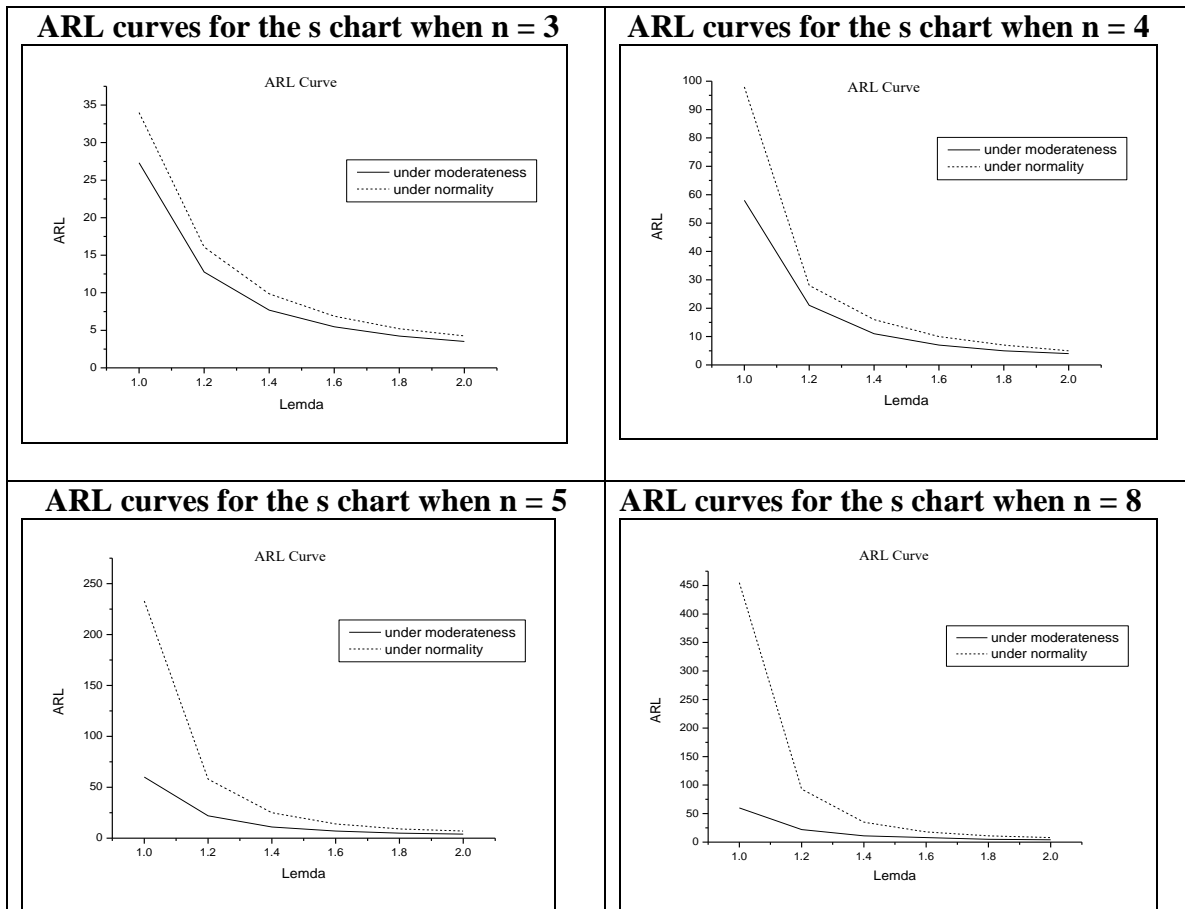


Figure 2:

From the figures 2, it is clear that for all values of n, even when there is up to 100% shift in the value of δ_0 (or σ_0), ARL under moderateness assumption are always smaller than ARL under normality assumption, which indicates that s chart under moderateness assumption is more effective than that of under normality assumption.

6. Summary

On the basis of OC curves analysis and ARL curves analysis, it is found that s chart under moderateness assumptions and having 3δ limits rather than 3σ limits are always more effective (perform better) than the charts under normality assumptions and having usual 3σ limits. So it is recommended that the control charts under moderateness assumption should be used for the best results.

7. Appendix

Moderate Distribution

Suppose the p.d.f. of a distribution of a random variable X is defined as,

$$f(x) = \frac{1}{\pi\delta} e^{-\frac{1}{\pi}\left(\frac{x-\mu}{\delta}\right)^2}, -\infty < X < \infty, \delta > 0$$

Then, the random variable X may be said to be following **moderate distribution** with parameters μ and δ and may be denoted as $X \sim M(\mu, \delta)$. It can be proved that,

- i. $\int_{-\infty}^{\infty} f(x) dx = 1$
- ii. Mean = $E(x) = \mu$
- iii. Mean deviation = $E[|X - \mu|] = \delta$
- iv. Standard deviation = $\sqrt{\frac{\pi}{2}}\delta$
- v. M.G.F = $M_x(t) = e^{\pi t + \frac{\pi}{4}\delta^2 t^2}$
- vi. $f(\mu-x) = f(\mu+x)$

It may be noted that the relationship between σ and δ is same as that in the normal distribution. Thus, the distribution of a random variable X having p.d.f. as defined above has location parameter as mean μ and scale parameter as mean deviation δ .

Just as the area of normal curve is measured in terms of σ from mean μ , the area of moderate distribution should be measured in terms of δ from mean μ . They have prepared moderate table pertaining to the area under standard moderate curve. From this table, it can be seen that

- a. $P(\mu-\delta < X < \mu+\delta) = 0.57506$ i.e. Mean \pm 1M.D covers 57.51% of area
- b. $P(\mu-2\delta < X < \mu+2\delta) = 0.88946$ i.e. Mean \pm 2M.D covers 88.95% of area.
- c. $P(\mu-3\delta < X < \mu+3\delta) = 0.98332$ i.e. Mean \pm 3M.D covers 98.33% of area.

References

- [1] D.C. Montgomery (2001): Introduction to statistical Quality Control (3rd Edition), John Wiley & Sons, Inc., New York,
- [2] E .L. Grant and R. S. Leavenworth (1988): Statistical Quality Control, Tata McGraw-Hill Publishing Company Limited, New Delhi
- [3] K.S.Tailor(2016): Moving Average And Moving Range Charts Under The Assumption Of Moderateness And Its 3δ Control Limits, Sankhya Vignana, December-2016, No.2, 18-31
- [4] K.S.Tailor(2017): Exponential Weighted Moving Average (EWMA) Charts Under The Assumption Of Moderateness And Its 3δ Control Limits, Mathematical Journal of Interdisciplinary Sciences, paper accepted for March 2017 issue
- [5] Naik V.D and Desai J.M. (2015): Moderate Distribution: A modified normal distribution which has Mean as location parameter and Mean Deviation as scale parameter, VNSGU Journal of Science and Technology, Vol.4, No. 1256-270
- [6] Naik V.D and Tailor K.S. (2015): On performance of \bar{X} and R-charts under the assumption of moderateness rather than normality and with 3δ control limits rather than 3σ control limits, VNSGU Journal of Science and Technology, Vol.4, No. 1, 243-255

- [7] V.D. Naik. and J.M. Desai.(2009): A Review of Debate on Mean Deviation Versus Standard Deviation and a Comparative Study of Four Estimators (Including Two Proposed Unbiased Estimators) of Mean Deviation of Normal Population, VNSGU Journal of Science &Technology, Vol. I, Issue-II, 220-229.
- [8] V.D. Naik. and J.M. Desai.(2010): Some more reasons for preferring Mean Deviation over Standard Deviation for drawing inference about first degree dispersion in numerical data, VNSGU Journal of Science &Technology, VNSGU Journal of Science &Technology, Vol. 2, December-2010
- [9] V.D. Naik and J.M. Desai (2015): Moderate Distribution: A modified normal distribution which has Mean as location parameter and Mean Deviation as scale parameter, VNSGU Journal of Science and Technology, Vol.4, No. 1, 256-270

*Corresponding author.

E-mail address: kalpesh_tlr@yahoo.co.in