Performance of Conventional $\overline{X}$ Chart for Auto-Correlated Data Using Subgroup Size of Ten

D. R. Prajapati

Assistant Professor, Department of Mechanical Engineering, PEC University of Technology (formerly Punjab Engineering College), Chandigarh, India

ABSTRACT— Control charts have been used to monitor and control the production process and generate some sort of signal when there are indications to the effect that the process has gone out of control. Conventional (Shewhart) $\overline{X}$ charts were originally developed by Walter F. Shewhart (1931). The limits of the chart are known as upper control limit (UCL) and lower control limit (LCL). He suggested two control limits i.e. UCL = $\overline{X}$ + 3σ/√n and LCL = $\overline{X}$ - 3σ/√n, where $\overline{X}$ is the target mean, σ is population standard deviation and n is the sample size. In implementing the $\overline{X}$ chart, it is assumed that the process outputs must be IID but usually there is some correlation among the data. When this correlation builds up automatically in the entire process, this phenomenon is called autocorrelation. It is found that the performance of conventional $\overline{X}$ chart deteriorates on increasing the level of correlation**.**

KEYWORDS— Average run length, sample size, $\overline{X}$ charts, IID data and Autocorrelation

I.INTRODUCTION

The process is assumed to be out of control when the sample average falls beyond these limits. The Shewhart $\overline{X}$ chart is more useful to detect large shifts (> 2σ) in process mean.

Several researchers examined the performance of control charts in the presence of autocorrelation. If process measurements are autocorrelated, then standard constructions of control charts will violate the assumption that samples are independent. The Type-I error rate for control charts is sensitive to autocorrelated data i.e. control charts are subjected to increased false alarms. Even at low levels of correlation, dramatic disturbances can occur in the chart properties.

A. Classification of Control charts

Basically, control charts can be classified as:

(i) Control charts for variables and

(ii) Control charts for attributes

The control charts for variables are used for quality characteristics which are measurable like length, diameter, tensile strength, temperature, viscosity etc. while the control charts for attributes are used for quality characteristics which are non-measurable. Control charts for variables may further be classified as:

- Control charts for monitoring the process means (averages)
- Control charts for monitoring the process dispersion
- Combined control charts
- Control charts for monitoring autocorrelated data

The process is assumed to be out of control when the sample average falls beyond these limits. If a variable has a normal distribution with mean of zero and standard deviation of one, and a $\overline{X}$ chart is used with control limits at ±3σ, the ARL will depend upon the sample size. ARLs for sample size of four are given in Table 1.

Amount of shift	ARL (n = 4)
0.0	370.3
1.0	44.0
2.0	6.3

Table 1 ARLs of $\overline{X}$ chart for sample size of four

A quality engineer may like his chart to generate a signal when the process average shifts by an amount equal to the process standard deviation. However an ARL of 44 may be too high as, on the average, the process will operate in an uncontrolled manner for a long period of time till 44 samples are inspected.

Performance of Conventional X-bar chart for Auto-correlated Data using subgroup size of ten

A. Autocorrelation

In general, it is assumed that the process outputs must be IID but usually there is some correlation among the data. When this correlation builds up automatically in the entire process, this phenomenon is called autocorrelation. The observations from the process output are usually positively correlated in most of the cases. In this case, if the current observation is on one side of the mean, the next observation will most likely be found on the same side of the mean. Positively correlated data are characterized by runs above and below the mean. Positive correlation is more often encountered in practice than negative autocorrelation.

When $\overline{X}$ charts are applied to the autocorrelated data, the false alarm rate increases and performance of the $\overline{X}$ chart is suspected. So improvement in the (Shewhart) $\overline{X}$ chart is needed to improve its performance for the correlated data. The following section deals with the literature review in this area.

Various researchers contributed in detecting the autocorrelation in data and modified the $\overline{X}$ charts. For the simulation of data, Markov chain, Monte Carlo, Time– series approach, MATLAB are widely used by the researchers. The research work done so far in modifying the $\overline{X}$ chart for the autocorrelated data is discussed as follows:

Kramer and Schmid (2000) discussed the behavior of the Shewhart residual chart and the modified Shewhart chart for unknown parameters of the underlying process. They focused on the estimation of the variance. For autoregressive models (AR models), they considered the estimation of the AR coefficients. They finally compared the ARL of the control chart with estimated parameters with the ARL of the scheme for known parameters and with the ARL for independent variables. Chiu and Yang (2001) suggested that neural networks can be used to identify shifts in correlated process parameters. They developed and trained the back propagation neural networks to identify shifts in process parameter values. They used all-possible-regression selection procedures for finding the appropriate numbers of input nodes to use in a neural network model. They found that the neural networks are successful at identifying shifts at one, one and half, and two standard deviations from the target. They concluded that the neural networks outperform the traditional SPC approaches in identifying shifts in process parameters.

Noorossana et al. (2003) presented an artificial neural network (ANN) model for detecting and classifying three types of non-random disturbances referred to as level shift, additive outlier and innovational outlier, which are common in autocorrelated processes. An autoregressive model of order one, AR(1) model was considered to characterize the quality characteristic of interest in a continuous process where autocorrelated observations were generated over time. The performance of the proposed procedure was evaluated through the use of a numerical example. Preliminary results indicated that the procedure could be used effectively to detect and classify unusual shocks in autocorrelated processes. Castagliola and Tsung (2005) investigated the effect of skewness on conventional autocorrelated SPC techniques and provided an effective approach based on a scaled weighted variance approach to improve SPC performance in such an environment. Gauri and Chakraborty (2008) used artificial neural network (ANN) to recognize different control chart patterns (CCPs). ANN with features extracted from the process data as input vector representation could facilitate efficient pattern recognition with a smaller network size. A set of seven shape features were selected, whose magnitudes are independent of the process mean and standard deviation under a special representation of the sampling interval in the control chart plot. Chen and Cheng (2009) found that in the design of the $\overline{X}$ control chart, both the sample size, 'm' of $\overline{X}$ and the control-limits factor, 'k' (the number of standard deviations from the center line) must be determined. They analyzed under the assumption that the quality characteristic, follows an autocorrelated process with known covariance structure but unknown marginal distribution shape.

Prajapati and Mahapatra (2009) provided a survey and brief summary of the work on the control charts for variables to monitor the process mean and dispersion from 1931 to 2008. Kiran et al. (2010) carried out a study on training algorithms for control charts pattern recognition and selected the best two patterns with their optimal structure for both type I and type II errors for generalization without early stopping and with early stopping and proposed the best one. Wu and Yu (2010) proposed a neural network-based identification model for both mean and variance shifts in correlated processes. The proposed model uses a selective network ensemble approach named DPSOEN (Discrete Particle Swarm Optimization) to obtain the improved generalization performance, which outperforms those of single neural network. Costa and Castagliola (2011) found that measurement error and autocorrelation have an adverse effect on the performance of $\overline{X}$ chart. In order to counteract the undesired effect of autocorrelation, they built-up the samples with non-neighboring items, according to the time the items are produced. They also measured the quality characteristic of each item of the sample several times. The chart's performance was assessed when multiple measurements are applied and the samples are built by taking one item from the production line and skipping one, two or more before selecting the next. McCracken and Chakraborti (2013) presented an overview of the literature covering some of the one-and two-chart schemes, including those that are appropriate in parameters known (standards known) and unknown (standards unknown) situations. They also discussed some of the joint monitoring schemes for multivariate processes, autocorrelated data, and individual observations. In addition, they discussed some available nonparametric schemes for jointly monitoring location and scale and recommended some areas of further research.

The formulation of the autocorrelated numbers is discussed in the next section.

III. FORMULATION OF AUTOCORRELATED DATA

In the autocorrelated series of observation, each individual observation is dependent upon the previous observation. A series of positively autocorrelated numbers with a mean of zero and standard deviation of one is generated, using the MATLAB at various levels of correlation (Φ). Assuming N pairs of observations on two variables, x and y. The correlation coefficient between x and y is given by equation (1). Some authors use coefficient of correlation (Φ) instead of "r".

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum (x_i - \bar{x})^2\right]^{1/2}\left[\sum (y_i - \bar{y})^2\right]^{1/2}} \qquad (1)$$

Where, the summations are over the N observations.

The series generated are positively correlated in nature. For each level of correlation (Φ), various schemes of $\bar{X}$ chart are developed, using MATLAB. The detailed procedure to implement the $\bar{X}$ chart is explained in the following section.

IV. PROCEDURE

The data are simulated with the help of MATLAB software. The autocorrelated numbers are generated from the series of data, taken from the standard normal distribution with mean of zero and standard deviation of one. The following procedure may be adopted to calculate the ARLs of the $\bar{X}$ chart.

Step 1 Take the observations from industry at random basis.

Step 2 Observations are generated randomly at a given mean and standard deviation.

Step 3 For simulation, 10,000 observations with a sample sizes (n) of 10 are generated.

Step 4 The observations are generated in such a way that there should be positive correlation with their previous data.

Step 5 Those sets of parameters of the $\bar{X}$ chart, which give the in-control ARL

Of approximately 370, are considered for comparison.

Step 6 For the selected combinations, the ARLs are calculated at various shifts in process mean at different width of the control limits (L) and at the each levels of correlation (Φ).

Step 7 Optimal schemes of modified $\bar{X}$ chart are obtained for levels of correlation (Φ) of 0.00, 0.25, 0.50, 0.75 and 1.00.

Next section deals with the computation of ARLs of the $\bar{X}$ chart for sample size of 10 and at different levels of correlation (Φ).

V. CONVENTIONAL $\bar{X}$ CHART FOR SAMPLE SIZE (N) OF TEN

Various schemes of the conventional $\bar{X}$ chart for sample size of ten are developed at various levels of correlation (Φ).

Table 2 shows the various schemes of conventional $\bar{X}$ chart for sample size (n) of ten at the levels of correlation (Φ) of 0.00 and 0.25.

Table 2 Schemes of Conventional $\bar{X}$ chart for sample size (n) of ten at Φ = 0.00 and 0.25

Shift (in mean)	Φ = 0.00		Φ = 0.25	
	L= 3.0	L=3.05	L= 3.0	L=3.05
0.00	370.4	370.4	370.4	370.4
0.25	212.5	242.5	222.2	312.5
0.50	116.3	136.7	125.5	166.7
0.75	66.3	76.3	78.7	105.3
1.00	36.7	39.0	39.2	49.0
1.25	19.6	25.8	22.2	28.8
1.50	10.0	12.8	11.3	16.8
1.75	8.1	9.1	8.4	11.1
2.00	5.7	6.0	6.0	7.0
2.50	2.8	3.2	3.0	3.5
3.00	1.7	1.7	1.8	2.1
3.50	1.4	1.3	1.4	1.4
4.00	1.0	1.2	1.0	1.3

Table 3 shows the various schemes of conventional $\bar{X}$ chart for sample size (n) of ten at the levels of correlation (Φ) of 0.50 and 0.75.

Table 3 Schemes of Conventional $\bar{X}$ chart for sample size (n) of ten at Φ = 0.50 and 0.

Table 4 shows the various schemes of Conventional $\bar{X}$ chart for sample size (n) of ten at the level of correlation (Φ) of 1.00.

M.R. Thansekhar and N. Balaji (Eds.): ICIET'14

Table 4 Schemes of Conventional chart for sample size (n) of ten at $\Phi = 1.00$

Shift (in mean)	$\Phi = 1.00$		
	L=3.0	L= 3.05	L= 3.10
0.00	370.4	370.4	370.4
0.25	277.8	333.3	336.2
0.50	212.5	263.2	272.3
0.75	132.6	178.6	182.5
1.00	77.5	95.2	98.3
1.25	45.5	68.0	71.6
1.50	25.6	34.5	46.0
1.75	22.0	30.1	32.2
2.00	15.7	20.1	22.8
2.50	8.4	10.4	11.2
3.00	4.3	5.73	5.9
3.50	3.0	3.7	3.7
4.00	2.1	2.5	2.4

Shift (in mean)	$\Phi = 0.00$	$\Phi = 0.25$	$\Phi = 0.50$	$\Phi = 0.75$	$\Phi = 1.00$
	L=3.0	L = 3.0	L=3.0	L= 3.0	L= 3.0
0.00	370.4	370.4	370.4	370.4	370.4
0.25	212.5	222.2	244.5	255.5	277.8
0.50	116.3	125.5	144.5	166.7	212.5
0.75	66.3	78.7	85.5	91.0	132.6
1.00	36.7	39.2	45.5	59.5	77.5
1.25	19.6	22.2	24.2	38.9	45.5
1.50	10.0	11.3	15.0	20.0	25.6
1.75	8.1	8.4	9.2	12.8	22.0
2.00	5.7	6.0	6.4	9.0	15.7
2.50	2.8	3.0	3.6	5.3	8.4
3.00	1.7	1.8	1.9	2.7	4.3
3.50	1.4	1.4	1.5	1.6	3.0
4.00	1.0	1.0	1.2	1.2	2.1

Table 5 Optimal schemes of $\overline{X}$ chart for sample size of ten (n = 10) at various levels of correlation (Φ)

These schemes of the conventional $\overline{X}$ chart for sample sizes of 10 are analyzed at various levels of correlation (Φ). Analysis of these schemes is presented in the next section.

VI. ANALYSIS

Those schemes of conventional $\overline{X}$ chart which give quick response to the shift in the process mean are known as optimal schemes for a given sample size (n) and level of correlation (Φ). These schemes have the lower ARLs than others schemes for all the shifts. The performance of these schemes at the various shifts in the process mean and at different levels of correlation (Φ) is given in the Tables 5.

Tables 2, 3, 4 and 5 represent the following facts:

(i) The in-control Average run length (ARL) of approximately 370 is maintained for all the suggested schemes of the conventional $\overline{X}$ chart.

(ii) For a particular sample size, when autocorrelation in the observations increases, the false alarm rate also increases.

(iii) The sensitivity of a chart to detect shift in the process mean depends upon the sample size.

(iv) The performance of conventional $\overline{X}$ chart deteriorates on increasing the level of correlation.

(v) At level of correlation (Φ) of 0.25 and for 1σ shift in process mean, the out of control ARL is 39.2 while it is 77.5 at $\Phi = 1.00$.

VII. CONCLUSIONS

The effect of autocorrelation on the sample size of ten for conventional (Shewhart) $\overline{X}$ chart for the correlated

M.R. Thansekhar and N. Balaji (Eds.): ICIET'14

observations is studied. The performance of the $\overline{X}$ chart is measured in terms of the Average Run Lengths (ARLs). It is concluded that if faster detection in the process mean is desired, it is recommended to use larger sample sizes (n). The optimal schemes for sample size of 10 at different levels of the correlation have been suggested that may be very useful from the quality control point of view for many industries. This control chart may also be very useful for the service sectors also.

REFERENCES

[1] H. Kramer and W. Schmid, "The influence of parameter estimation on the ARL of Shewhart type charts for time series", *Statistical Papers*, Volume 41, Number 2, pp. 173-196, 2000.

[2] C. C. Chiu and M. H. Yang, "Shifts recognition in correlated manufacturing process parameters using neural network", *International Journal of the Chinese Institute of Industrial Engineers*, Volume 18, Issue 3, pp. 86 – 94, 2001.

[3] R. Noorossana, M. Farrokhi and A. Saghaei, "Using neural networks to detect and classify out-of-control signals in autocorrelated processes", *Quality and Reliability Engineering International*, Volume 19, Issue 6, pp. 49 –504, 2003.

[4] P. Castagliola and F. Tsung, "Autocorrelated SPC for non-normal situations", *Quality and Reliability Engineering International*, Volume 21, Issue 2, pp. 131-161,2005.

[5] S. K. GAURI, AND S. CHAKRABORTY, "IMPROVED RECOGNITION OF CONTROL CHART PATTERNS USING ARTIFICIAL NEURAL NETWORKS", *INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY*, VOLUME 36, NUMBER 11-12, PP. 1191-1201,2008.

[6] H. CHEN, AND Y. CHENG, "DESIGNING $\overline{X}$ CHARTS FOR KNOWN AUTOCORRELATIONS AND UNKNOWN MARGINAL DISTRIBUTION", *EUROPEAN JOURNAL OF OPERATIONAL RESEARCH,* VOLUME 198, NUMBER 2, PP. 520 – 529, 2009.

[7] D. R. Prajapati and P. B. Mahapatra, "Control charts for variables to monitor the process mean and dispersion: a literature review", *International Journal of Productivity and Quality Management*, Volume 4, Number 4, pp. 476 – 520, 2009.

[8] B. Wu and J. B. Yu, "A neural network ensemble model for on-line monitoring of process mean and variance shifts in correlated processes", *Expert Systems with Applications*, Volume 37, Issue 6, pp. 4058-4065, 2010.

[9] A. F. B. Costa, and P. Castagliola, (2011), "Effect of measurement error and autocorrelation on the $\overline{X}$ chart", *Journal of Applied Statistics,* Volume 38, Issue 4, pp. 661 – 673, 2011.

[10] A.K. McCracken and S. Chakraborti, "Control Charts for Joint Monitoring of Mean and Variance: An Overview", Quality Technology of Quantitative Management, Vol. 10, No. 1, pp. 17-36, 2013.

BIOGRAPHY

Dr. D. R. PRAJAPATI

Working as an Assistant Professor in the department of Mechanical Engineering, PEC University of Technology (formerly Punjab Engineering College), Chandigarh-160012 (India).

He is having the teaching and research experience of more than 18 years and published more than 75 research papers in international and national journals of repute and in the proceedings of the conferences. He is also reviewer of 6 international journals. He also guided 2 Ph.D. and more than 15 post graduate theses and guiding 4 research scholars at present. He has also chaired international and national conference in India and abroad. He also organized two short term courses and one All India Conference for the faculty of technical institutions and industries. He is also recipient of first D. N. Trikha research award for excellent research publications in international journal for the year 2009 in PEC University of Technology.

M.R. Thansekhar and N. Balaji (Eds.): ICIET'14