Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 8, Issue 1 (2015): 25-30

THE POWER OF $\overline{X}$-CHART IN PRESENCE OF DATA CORRELATION

Ab Latif Dar* and J. R. Singh
School of Studies in Statistics
Vikram University, Ujjain-456010, M.P. India
E Mail: *lateefdar.2007@rediffmail.com

Received December 31, 2014
Modified April 01, 2015
Accepted April 15, 2015

Abstract

A fundamental assumption in the development of $\overline{X}$ control chart is that the underlying distribution of the quality characteristic is normal and the process data is normally distributed. However we may have certain situations where the assumption of independence and normality are either fully or partially violated. In this paper an attempt has been made to examine the power of $\overline{X}$ chart when the assumption of independence of the data is violated. Expressions for the power of $\overline{X}$-chart are derived for different values of correlation coefficient.

Key Words: Control Chart, Correlation, Power of $\overline{X}$-Chart.

1. Introduction

Control charts are widely used tools of statistical quality control in industrial environments since its inception by Shewart in 1920's. However in recent years, there has been a growing criticism of its application in certain types of manufacturing wherein data correlation is prevalent. Shewart control chart is based on the assumptions that the process data are independent and normally distributed. Out of these assumptions, it is very sensitive to the data independent assumption. This is often violated in data from most of the process industries and even in certain types of automated manufacturing. This is due to the fact that the increase of automation in manufacturing process/ process industries makes it possible to collect each and every observation through automatic measuring equipment and the data points, thus obtained are auto–correlated, since the frequency of observation is less than the process time constant Montgomery (1991). Further, the data correlation has also been observed in data from the conglomerated manufacturing, wherein the characteristics of both manufacturing and process industries are observed Box et al., (1970). Control charts are useful tools for monitoring/controlling a manufacturing process. With properly chosen control limits, a control chart can detect a shift from a "good" quality to a "bad" one. When the measurement denoted by X,of a particular characteristic of a product is used to gauge the quality of the product , the most commonly used charts are the X chart, the $\overline{X}$ chart or (Shewar chart) and cumulative sum (CUSUM) chart . These charts are easy to construct, visualize and interpret and most important and have been proven effective in practice. Shewart control charts are widely used in industry, usually with the assumption that the data follow the normal distribution and are independent. However, this assumption is not fulfilled for many data sets. If it is assumed that the random variables are independent, when infact they are serially correlated, it has been shown by

Montgomery and Mastrangelo (1991), Margah and Woodall (1992), Wordell et al. (1994) that the classical 3σ- limits in Shewart charts for monitoring the process mean are not suitable. Neuhardt (1987) studied the effect of correlation within subgroup on control charts. Yang and Hancock (1990) extended Neuhardt's work to conduct simulation studies to determine the effect of correlated data on $\overline{X}$, R, S and S^2 charts . From their studies, Yang and Hancock (1990) concluded that if a positive correlation exists but is not recognized , the actual type I error probability of $\overline{X}$ chart will be significantly larger than assumed; however, failure to recognize the correlation will slightly increase the type I error probabilities in R, S and S^2 charts do not need to be revised even if correlation exists. For the assumption of normality, if the measurements are really normally distributed, the statistic $\overline{X}$ is also normally distributed. Rossoul et.al. (2010) studied the effect of correlation on linear profile monitoring. where as Padgett et al. (1992) have studied the α-risk of $\overline{X}$ control chart. Kanazuka (1986) studied the effect of measurement error on the power of $\overline{X}$ Chart. Singh and Singh (1982) relaxed the assumption of independence in the classical quality control model and replaced it by dependence introduced via second order auto regressive model. The above studies show that auto-correlation has a significant effect on the performance of standard control charts. Singh and Kulkarni (2009) studied the effect of Yule's process on the power of $\overline{X}$ chart and assumed the observations to be independent. In this paper an attempt has been made to study the power of $\overline{X}$ chart, when the observations are correlated.

2. Power of $\overline{X}$ Chart in presence of data correlation

In this development it is assumed that the process has a normal distribution with mean μ and variance σ^2. It is further assumed that at the time of determining the control limits the process is in a statistical control, and the same device is used as will be employed for latter measurements. We further assume that the observations came from the normal population and the observations are correlated or we can say that the assumption of independence is violated in this case. Thus the data used for establishing the limits on the control chart comes from a process that is N(μ, σ^2). When the process shifts, the data is assumed to come from a N(μ', σ^2T^2) population. If samples of size n are taken from the population N(μ', σ^2T^2) and the value of $\overline{X}$ is plotted with control limits of $\mu \pm 3\sqrt{\sigma^2/n}$, the power of detecting the change of process is given by the following formula:

$$P_{\bar{X}} = \Pr\{ \overline{X} \geq \mu + 3\sqrt{\sigma^2/n}\} + \Pr\{ \overline{X} \leq \mu - 3\sqrt{\sigma^2/n}\} \tag{2.1}$$

Converting to a standard normal distribution, we have

$$Z=(\overline{X} - \mu') / \sqrt{\sigma^2T^2/n} \tag{2.2}$$

Where $T^2 = 1 + (n-1)\rho$

As we assume that the data comes from the normal population and the observations are dependent. Thus the variance of $\overline{X}$ under the correlated data is given by

$$V(\overline{X}) = \frac{\sigma^2}{n}\{1 + (n-1)\rho\}$$
$$= \frac{\sigma^2}{n} T^2$$

where $T^2 = \{1 + (n-1)\rho\}$,where σ^2 is assumed to be known, ρ is the correlation coefficient and n is the sample size.

Now assuming the new variable, equation (2.1) can be written as:

$$P_{\bar{X}} = \Pr\{Z \geq (\mu - \mu')\sqrt{\frac{n}{\sigma^2 T^2}} + \frac{3}{T}\} + \Pr\{Z \leq (\mu - \mu')\sqrt{\frac{n}{\sigma^2 T^2}} - \frac{3}{T}\}$$

$$= \Pr\{Z \geq -(\mu' - \mu)\sqrt{\frac{n}{\sigma^2 T^2}} + \frac{3}{T}\} + \Pr\{Z \leq -(\mu' - \mu)\sqrt{\frac{n}{\sigma^2 T^2}} - \frac{3}{T}\}$$

$$= \Pr\{Z \geq (\frac{-d\sqrt{n}}{T} + \frac{3}{T})\} + \Pr\{Z \leq (\frac{-d\sqrt{n}}{T} - \frac{3}{T})\}$$

$$= \Pr\{Z \geq \frac{1}{T}(-d\sqrt{n} + 3)\} + \Pr\{Z \leq \frac{1}{T}(-d\sqrt{n} - 3)\}$$

$$= \Pr\{Z \geq \frac{1}{T}(3 - d\sqrt{n})\} + \Pr\{Z \leq \frac{1}{T}(-3 - d\sqrt{n})\}$$

$$= \Phi\{\frac{1}{T}(-3 + d\sqrt{n})\} + \Phi\{\frac{1}{T}(-3 - d\sqrt{n})\}$$

where $(\mu' - \mu)/\sigma = d$, is the change of process average and

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Z} \exp(\frac{u^2}{2})$$

3. Numerical illustrations and results

For the purpose of illustrating the effect of correlation on the power of $\bar{X}$ chart, we have determined the values of power function $P_{\bar{X}}$ for independent observations (ρ=0) and for different values of correlation coefficient ρ. The values of power function have been calculated and the results are presented in table 3.1 for n =5. In order to give visual comparison of the power functions for different values of ρ, a curve have been drawn and shown in fig. 3.2 , which illustrates the relationship between the change of process average d and the power of detecting this change $P_{\bar{X}}$ when n =5 . The power depends upon the magnitude of process change with correlated observations. From the table 3.1 and fig. 3.2 it is evident that the power of $\bar{X}$ chart increases as the process average d increases and decreases with increase in the value of correlation coefficient as $P_{\bar{X}}$ = 0.00338, 0.00307, 0.00296, 0.00290, 0.00286 and 0.00283 for ρ = 0, 0.2, 0.4, 0.6, 0.8 and 1(d=0.1).The power of $\bar{X}$ chart is maximum for ρ = 0. It is also evident from the table 3.1 and fig. 3.2 that the power of $\bar{X}$ chart increases as the sample size increases. So on designing control charts for mean it is suggested that utmost care must be taken while dealing with the correlated data and there is need to investigate more deeply the performance of $\bar{X}$ control chart in presence of data correlation.

d	ρ					
	0	0.2	0.4	0.6	0.8	1.0
0.1	0.00338	0.00307	0.00296	0.00290	0.00286	0.00283
0.2	0.00563	0.00426	0.00376	0.00350	0.00335	0.00324
0.3	0.01005	0.00644	0.00520	0.00458	0.00420	0.00395
0.4	0.01767	0.00994	0.00743	0.00620	0.00548	0.00500
0.5	0.02994	0.01519	0.01065	0.00850	0.00725	0.00644
0.6	0.04863	0.02278	0.01515	0.01163	0.00963	0.00836
0.7	0.07568	0.03339	0.02125	0.01579	0.01275	0.01083
0.8	0.11292	0.04780	0.02936	0.02122	0.01676	0.01398
0.9	0.16169	0.06682	0.03991	0.02818	0.02183	0.01791
1.0	0.22245	0.09122	0.05336	0.03696	0.02816	0.02278
1.1	0.29449	0.12168	0.07017	0.04786	0.03596	0.02874

1.2	0.37573	0.15867	0.09080	0.06120	0.04546	0.03595
1.3	0.46291	0.20235	0.11562	0.07730	0.05688	0.04458
1.4	0.55191	0.25252	0.14492	0.09642	0.07046	0.05481
1.5	0.63837	0.30856	0.17884	0.11881	0.08640	0.06682
1.6	0.71827	0.36947	0.21736	0.14464	0.10489	0.08077
1.7	0.78853	0.43385	0.26029	0.17401	0.12609	0.09681
1.8	0.84730	0.50004	0.30721	0.20692	0.15011	0.11508
1.9	0.89408	0.56622	0.35752	0.24325	0.17701	0.13568
2.0	0.92951	0.63060	0.41044	0.28277	0.20676	0.15867
2.1	0.95503	0.69150	0.46505	0.32514	0.23928	0.18408
2.2	0.97253	0.74754	0.52032	0.36991	0.27442	0.21187
2.3	0.98394	0.79770	0.57520	0.41651	0.31193	0.24199
2.4	0.99102	0.84137	0.62866	0.46432	0.35150	0.27428
2.5	0.99520	0.87835	0.67975	0.51265	0.39275	0.30856
2.6	0.99755	0.90881	0.72763	0.56079	0.43524	0.34461
2.7	0.99881	0.93321	0.77167	0.60805	0.47850	0.38212
2.8	0.99944	0.95222	0.81139	0.65377	0.52200	0.42077
2.9	0.99975	0.96663	0.84654	0.69735	0.56525	0.46021
3.0	0.99990	0.97726	0.87705	0.73828	0.60773	0.50004
3.1	0.99996	0.98488	0.90304	0.77617	0.64896	0.53987
3.2	0.99998	0.99019	0.92475	0.81073	0.68851	0.57930
3.3	0.99999	0.99379	0.94254	0.84179	0.72600	0.61795
3.4	1.00000	0.99617	0.95684	0.86931	0.76111	0.65546
3.5	1.00000	0.99770	0.96811	0.89332	0.79360	0.69150
3.6	1.00000	0.99865	0.97684	0.91398	0.82332	0.72578
3.7	1.00000	0.99923	0.98346	0.93148	0.85018	0.75807
3.8	1.00000	0.99957	0.98839	0.94610	0.87417	0.78818
3.9	1.00000	0.99977	0.99199	0.95814	0.89534	0.81597
4.0	1.00000	0.99988	0.99457	0.96790	0.91380	0.84137
4.1	1.00000	0.99994	0.99638	0.97570	0.92971	0.86436
4.2	1.00000	0.99997	0.99763	0.98184	0.94326	0.88496
4.3	1.00000	0.99998	0.99848	0.98660	0.95467	0.90322
4.4	1.00000	0.99999	0.99904	0.99025	0.96415	0.91926
4.5	1.00000	1.00000	0.99940	0.99300	0.97194	0.93321
4.6	1.00000	1.00000	0.99964	0.99504	0.97827	0.94522
4.7	1.00000	1.00000	0.99978	0.99653	0.98335	0.95545
4.8	1.00000	1.00000	0.99987	0.99761	0.98737	0.96408
4.9	1.00000	1.00000	0.99993	0.99837	0.99053	0.97129
5.0	1.00000	1.00000	0.99996	0.99891	0.99297	0.97726

Table 3.1: Value of Power Function for n=5

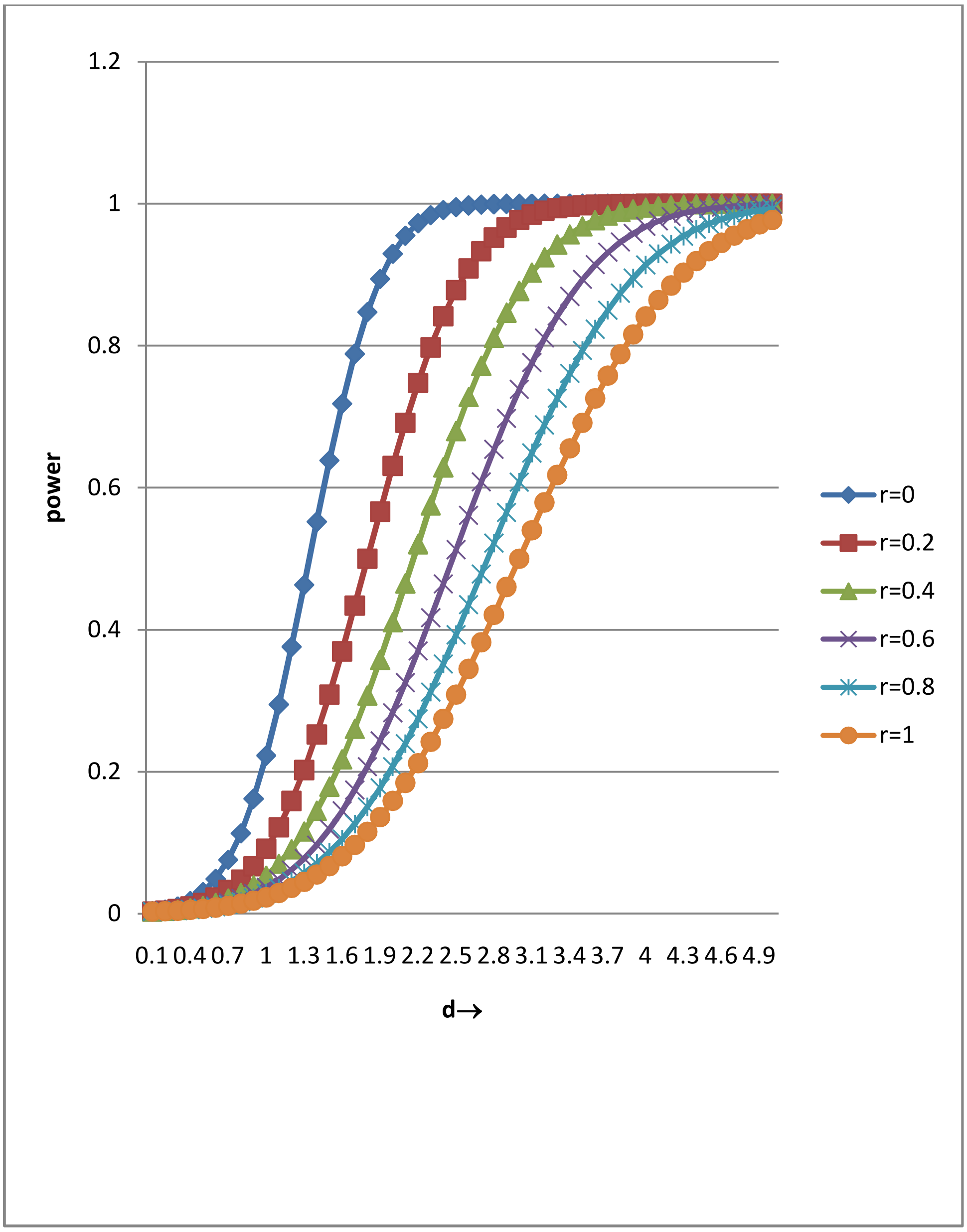

Fig. 3.2: Power curve for different values of ρ and n = 5

References

1. Box, G.E.P., Jenkins, G.M. and Rinsel, S.M. (1994). Time series analysis: Forecasting and control, 2nd ed. Holden-Day, San Francisco, CA.
2. Chou, C.Y. , Chen, C.H and Liu, H.R. (2005). Acceptance control charts for non-normal data, Journal of Applied Statistics, 32, p. 25–36.
3. Kanazura, T. (1986). The effect of measurement error on the power of $\bar{X} - R$ chart, J. of Quality Technology,18(2), p. 91-95.
4. Maragah,H.D, and Woodall, W.H.(1992). The effect of autocorrelation on the retrospective $\bar{X}$ chart, Journal of Statistical Computation and Simulation, 40, p. 29-42.
5. Montgomery, D.C. (1991). Introduction to Statistical Quality Control, 2nd ed. J. Wiley & Sons, New York, NY.
6. Montgomery, D.C., and Mastrangelo, C.M. (1991). Some statistical process control methods for auto-correlated data (with discussion), Journal of Quality Technology, 23, p. 179-204.
7. Neuhardt, J.B. (1987). Effect of correlated sub-samples in statistical process control, .IIE transactions, 19, p. 208-214.
8. Padgett, C. S., Tombs, L. A. and Padgett, W. J. (1992). On the α-risks for Shewart control, Comm. Statist. Simulation and Computation, 21, p. 1125-1147.
9. Rossoul,N., Abbas, S. and Mehdi,D. (2010). Linear profile monitoring in the presence of non-normality and autocorrelation, International journal of Industrial Engineering and Production Research, 21(4), p. 221-230.
10. Singh, H.R. and Singh, J. R. (1982). Variable Sampling Plan under second order Auto-regressive model. Indian Association of Products, Quality and Reliability Transactions , 7(2), p. 97-104.
11. Singh, J. R. and Kulkarni, K. (2009). The effect of Yule's process on the power of $\bar{X}$ chart, Varahmihir, J. Math. Sci., 9(1), p. 123-126..
12. Wardell,D.G., Moskowitz,H. and Plante, R.D.(1994). Run length distribution of special- cause control charts for correlated processes (with discussion), Technometrics, 36(1), p. 3-27.
13. Yang,K. and Hancock, W.M.(1990). Non-normality and design of control charts for averages, Decision Sciences, 23, 1099-1113.