Ultra Scientist Vol. 24(1)A, 227-234 (2012).

# The effect of yule's process on the power of $\overline{X}$-chart

J.R. SINGH* and K. KULKARNI**

**School of Studies in Statistics
Vikram University, Ujjain, M.P. (INDIA)
*jrsingh@gmail.com, **Ketki333@rediffmail.com

(Acceptance Date 2nd April, 2012)

## Abstract

An attempt has been made to study the effect of Yule's process on the Power of $\overline{X}$-chart. Process measurements which are use in construction of $\overline{X}$-chart involve several of errors. These include the inherent variability in the process and the error due to Yule's process.

*Key words:* Auto-Regressive Model of order two (*i.e.*,AR-2) or Yule's Model, $\overline{X}$-Control Chart, Correlation, Auto-Correlation.

## 1 Introduction

A basic assumption made in most traditional applications of control charts is that the observations from the process are independent. When the mean of observations is being monitored, the mean is assumed to be constant at the target value until a special cause occurs and produces a change in the mean. However, for many processes, such as in the chemical and process industries, there may be correlation between observations that are closely spaced in time. Correlation which is not a significant problem for an FSI chart may become problematic for a VSI chart because some of the observations will be taken using a relatively short sampling interval. The effect of correlated observations on the performance of FSI control charts has been studied by several authors.

Goldsmith and Whitfield[3]; Johnson and Bagshaw[5]; Bagshaw and Johnson[1]; Harris and Ross[4]; Yashchin[8] and VanBrackle and Reynolds[7] investigated the effects of correlation observation on CUSUM charts. Harris and Ross[4] and VanBrackle and Reynolds[7] also investigated the effects on EWMA charts. Vasilopoulos and Stamboulis (1978); Maragah and Woodall (1992) and Padgett *et al*. (1992) investigated $\overline{X}$ charts with correlated observations, where the control limits are estimated from the data[6].

The decision - making method is the same for both the first and second steps: if the points fall outside the corresponding control limits, the null hypothesis is rejected. Rejection of at least one of the simple null hypothesis leads to rejection of the general null hypothesis.

Thus, a positive decision as to whether a sampling point belongs to the universe is based on the logical and of the simple null hypothesis, $\sigma^2 = \sigma_0^2$ and $\mu = \mu_0$. To determine whether the process mean is at the standard value, the control chart for averages ($\overline{x}$) is most widely used, while charts based on either the sample range R or the sample standard deviation S are used to monitor process variability. A basic assumption made in most traditional applications of control charts is that the observations from the process are independent.

In some situations in which there is correlation between the observations, it may not be very realistic to assume that the process mean is constant until a special cause occurs. It may be more realistic to assume that the process mean is continually wandering even though no special cause can be identified. For some situations of this type, the objective of process monitoring may be the application of some kind of engineering feedback control method in addition to the detection of special causes (see, e.g., Macgregor (1990)). This paper, however, will not consider feedback control problems, but will concentrate on the detection of special causes in a process which is continually wandering about an overall mean. It is assumed that a special cause produces a shift in the overall mean.

In this paper an attempt has been made to study the effect of Yule's process on the Power of $\overline{x}$-chart. Process measurements which are use in construction of $\overline{x}$-chart involve several of errors. These include the inherent variability in the process and the error due to Yule's process.

2 *Model Description :*

For processes in which there is correlation between observations, a more reasonable model may be

$$x_t = \mu + \xi_t \qquad t = 1,2,...n, \quad (2.1)$$

where $\mu$ is the mean at time t. The assumption here is that the mean is not a fixed constant but rather continually wanders over time. As in the Durbin and Watson[2] d-statistic can be used to detect the presence or absence of serial correlation. The problem, however, is what to do once the suspicion of dependence via the auto/correlation test is confirmed. If autocorrelation exist we use identification techniques to define the nature of $\xi_t$. When identification is complete, the likelihood function can provide maximum likelihood estimate of the parameters of the identified model.

Suppose that a correlation test revealed the presence of data dependence and the identification technique suggested autoregressive model of order two (*i.e.* AR-2) or Yule's model, then we can express $\xi_t$ of equation (2.1) as

$$\xi_t = \alpha_1 \xi_{t-1} + \alpha_2 \xi_{t-2} + \varepsilon_t, \; t=1,2,...,n \quad (2.2)$$

where

(i) $\varepsilon_t \sim N(O, \sigma_\varepsilon^2)$ (2.3)

(ii) $Cov(\varepsilon_t, \varepsilon_n) = \begin{cases} \sigma_\varepsilon^2 & t = \tau \\ 0 & t \neq \tau \end{cases}$

For stationarity, the roots of the characteristic equation of the process

$$\phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 = 0 \qquad (2.4)$$

must lie outside the unit circle, which implies that the parameters $\alpha_1$ and $\alpha_2$ must lie in the triangular region, *i.e.*,

$$\alpha_2 + \alpha_1 < 1,\ \alpha_2 - \alpha_1 < 1 \text{ and } -1 < \alpha_2, < 1$$

Using $\phi(B)\ \rho_k = 0$ and Yule's process has an autocorrelation function given by the second order difference equation

$$\rho_k = \alpha_1\ \rho_{k-1} + \alpha_2\ \rho_{k-2} \qquad (2.5)$$

The variance of Yule's process is given by

$$\sigma^2 = \left(\frac{1-\alpha_2}{1+\alpha_2}\right) \frac{\sigma_\varepsilon^2}{\left[(1-\alpha_2)^2 - \alpha_1^2\right]} \qquad (2.6)$$

Now if $G_1^{-1}$ and $G_2^{-1}$ are the roots of the characteristic equation of the process given by equation (2.4) then

$$G_1 = \frac{\alpha_1 + (\alpha_1^2 + 4\alpha_2)^{1/2}}{2} \qquad (2.7)$$

$$G_2 = \frac{\alpha_1 - (\alpha_1^2 + 4\alpha_2)^{1/2}}{2} \qquad (2.8)$$

For stationary we require that $|Gi| < 1$. Thus, three situations can theoretically arise :

(i) Roots $G_1$ and $G_2$ are real and distinct (i.e. $\alpha_1^2 + 4\alpha_2 > 0$)

(ii) Roots $G_1$ and $G_2$ are real and equal (i.e., $\alpha_1^2 + 4\alpha_2 = 0$)

(iii) Roots $G_1$ and $G_2$ are complex conjugate (i.e., $\alpha_1^2 + 4\alpha_2 < 0$)

When the correlation is present in the data, we have for the distribution of the sample mean $\overline{x}$, its mean and variance given by

$$E(\overline{x}) = \mu$$

$$\mathrm{Var}(\overline{x}) = \frac{\sigma^2}{n} \lambda_{ap}(\alpha_1, \alpha_2, n)$$

$$= \frac{\sigma^2}{n} T^2 \qquad (2.9)$$

Where $T^2 = \lambda_{ap}(\alpha_1, \alpha_2, n)$ depends on the nature of the roots $G_1$ and $G_2$ and for different situations is given as follows :

(i) If $G_1$ and $G_2$ are real and distinct,

$$\lambda_{ap}(\alpha_1, \alpha_2, n) = \left[\frac{G_1(1-G_2^2)}{(G_1 - G_2)(1+G_1G_2)} \lambda(G_1, n) - \frac{G_2(1-G_1^2)}{(G_1 - G_2)(1+G_1G_2)} \lambda(G_2, n)\right]$$

$$= \lambda_{rd}(\alpha_1, \alpha_2, n), \qquad (2.10)$$

where $\lambda(G, n) = \left[\frac{1+G}{1-G} - \frac{2G}{n} \frac{(1-G^n)}{(1-G)^2}\right]$

(ii) If $G_1$ and $G_2$ are real and equal

$$\lambda_{ap}(\alpha_1, \alpha_2, n) = \left(\frac{1+G}{1-G}\right) - \frac{2G}{n} \frac{(1-G^n)}{(1-G)^2}$$

$$\left[1+\frac{(1+G)^2\,(1-G^n)-n\,(1-G^2)(1+G^2)}{(1+G^2)(1-G^n)}\right]^1$$

$$= \lambda_{re}(\alpha_1,\alpha_2,n), \tag{2.11}$$

(iii) If $G_1$ and $G_2$ are complex conjugate

$$\lambda_{ap}(\alpha_1,\alpha_2,n)=\left[Y(d,u)+\frac{2d}{n}(W(d,u,n)+z(d,u,n))\right]$$

$$= \lambda_{cc}(\alpha_1,\alpha_2,n), \tag{2.12}$$

Where

$$Y(d,u)=\frac{1-d^4+2d(1-d^2)\cos u}{(1+d^2)(1+d^2-2d\cos u)}$$

$$W(d,u,n)=\frac{2d(1+d^2)\sin u-(1+d^4)\sin 2u-d^{n+4}\sin(n-2)u}{(1+d^2)(1+d^2-2d\cos u)^2\sin u}$$

$$Z(d,u,n)=\frac{2d^{n+3}\sin(n-1)u-2d^{n+1}\sin(n+1)u+d^n\sin(n+2)u}{(1+d^2)(1+d^2-2d\cos u)^2\sin u}$$

$$d^2=-\alpha_2$$

and $$u=\cos^{-1}\left(\frac{\alpha_1}{2d}\right)$$

*3 Power of $\overline{X}$-chart for Yule's process :*

In this development it is assumed that the process has a normal distribution with mean $\mu$ and variance $\sigma^2$. It is assumed that at the time of determining the control limits the process is in a state of statistical control, and the same device is used as will be employed for later measurements. Thus, the data used for establishing the limits on the control charts comes from a process that is $N(\mu,\sigma^2)$. When the process shifts, the data is assumed to come from an $N(\mu', \sigma^2T^2)$ population with Yule's process. If the samples of size n are taken from the population $N(\mu', \sigma^2T^2)$ 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 :

$$P_{\bar{x}} = P_r\left\{\bar{x} \geq \mu + 3\sqrt{\sigma^2/n}\right\} + P_r\left\{\bar{x} \leq \mu - 3\sqrt{\sigma^2/n}\right\} \tag{3.1}$$

Converting to a standard normal distribution we have

$$Z = (\bar{x} - \mu')/\sqrt{\sigma^2 T^2/n} \tag{3.2}$$

Using the new variable, equation (3.1) can be expressed :

$$P_{\bar{x}} = P_r\left\{Z \geq (\mu - \mu^1)\sqrt{\frac{n}{\sigma^2 T^2}} + 3/T\right\} + P_r\left\{Z \leq (\mu - \mu^1)\sqrt{\frac{n}{\sigma^2 T^2}} - 3/T\right\}$$

$$= P_r\left\{Z \geq -\frac{d}{T}\sqrt{n} + \frac{3}{T}\right\} + P_r\left\{Z \leq -\frac{d}{T}\sqrt{n} - \frac{3}{T}\right\}$$

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

$$= \Phi\left\{\frac{1}{T}(-3 + d\sqrt{n})\right\} + \Phi\left\{\frac{1}{T}(-3 - d\sqrt{n})\right\} \tag{3.3}$$

where $(\mu'-\mu)/\sigma = d$, and

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

## 4 Numerical Illustration and Results

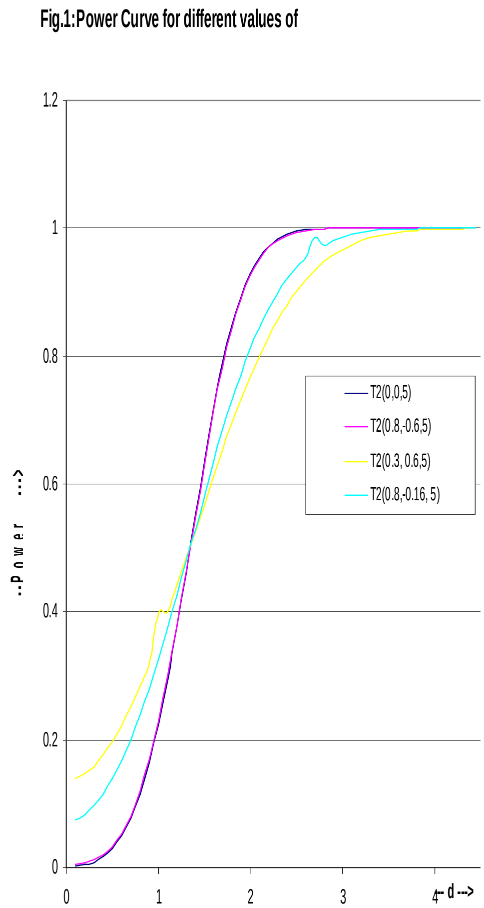
For the purpose of illustrating the effect of Yule's process, we have determined the values of power function P $\overline{X}$ for three different situations of the root from the equation (2.10), (2.11) and (2.12). The values of power function have been calculated and the results presented in Table 1 and Table 2 for n=5 and 7. For three different situations (i.e. (i) roots are read and distinct (ii) roots are real and equal and (iii) roots are complex conjugate). In order to give visual comparison of the power functions for three different situation and independent case, a curves have been drawn and shown in Fig. 1 and Fig. 2, which illustrates the relationship between the change of the process average d and the power of detecting this change P $\overline{X}$ when n=5 and 7. The power depends on the magnitude of the process change and three different situations of the roots with independent observation. When the roots are complex conjugate the power curves are approximately same as the observations are independent. The use of larger n will result in improvement for power function. The functions are seriously affected when the roots are real and distinct, real and equal. If a process generates data that are inherently Yule's process, due to natural process dynamics, power function should be used to monitor any engineering control system that has establish to minimize process variation about a mean target.

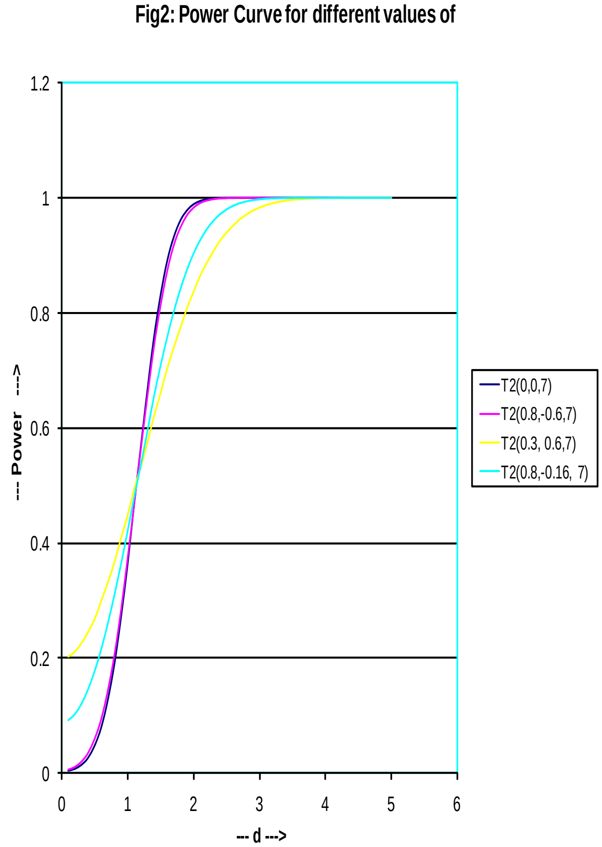
**Table 2.1 : Value of Power Function for Yule's Model and n = 5**

| d | $T^2(0,0,5)$ | $T^2(0.8,-0.6,5)$ | $T^2(0.3, 0.6,5)$ | $T^2(0.8,-0.16, 5)$ |
|---|---|---|---|---|
| 0.1 | 0.0036 | 0.0041 | 0.1386 | 0.0754 |
| 0.5 | 0.0311 | 0.0330 | 0.1964 | 0.1382 |
| 1.0 | 0.2238 | 0.2298 | 0.3990 | 0.3273 |
| 1.5 | 0.6368 | 0.6330 | 0.5682 | 0.5832 |
| 2.0 | 0.9292 | 0.9250 | 0.7674 | 0.8106 |
| 2.5 | 0.9952 | 0.9942 | 0.9015 | 0.9395 |
| 3.0 | 0.9998 | 0.9998 | 0.9672 | 0.9868 |
| 3.5 | 0.9999 | 0.9999 | 0.9919 | 0.9980 |
| 4.0 | 0.9999 | 0.9999 | 0.9985 | 0.9998 |
| 4.5 | 1.0000 | 1.0000 | 0.9998 | 0.9999 |
| 5.0 | 1.0000 | 1.0000 | 0.9999 | 0.9999 |

**Table 2.2: Value of Power Function for Yule's Model and n = 7**

| d | $T^2(0,0,7)$ | $T^2(0.8,-0.6,7)$ | $T^2(0.3, 0.6,7)$ | $T^2(0.8,-0.16, 7)$ |
|---|---|---|---|---|
| 0.1 | 0.0038 | 0.0071 | 0.2019 | 0.0928 |
| 0.5 | 0.0476 | 0.0607 | 0.2673 | 0.1783 |
| 1.0 | 0.3633 | 0.3746 | 0.4482 | 0.4215 |
| 1.5 | 0.8314 | 0.8132 | 0.6604 | 0.7089 |
| 2.0 | 0.9889 | 0.9829 | 0.8366 | 0.9032 |
| 2.5 | 0.9998 | 0.9995 | 0.9395 | 0.9799 |
| 3.0 | 0.9999 | 0.9999 | 0.983 | 0.9974 |
| 3.5 | 1.0000 | 0.9999 | 0.9965 | 0.9998 |
| 4.0 | 1.0000 | 1.0000 | 0.9995 | 0.9999 |
| 4.5 | 1.0000 | 1.0000 | 0.9999 | 0.9999 |
| 5.0 | 1.0000 | 1.0000 | 0.9999 | 0.9999 |




The traditional control charts should not be applied without modification to Yule's process. An appropriate application of power function in this case is to monitor the engineering control system applying traditional techniques to the residual or deviation from target. In practical situations, however, the case of real and equal roots hardly arises. Moreover, this is too simple a case for calculations and included only for the sake of completion. The power curves looks different for other gradients and there are cases where the traditional approach is more powerful.

## References

1. Bagshaw, M. and Johnson, R.A., The Effect of Serial Correlation on the Performance of CUSUM Tests II. *Technometrics* 17, pp. 73-80 (1975).
2. Durbin, J. and Watson, G. S., Testing for Serial Correlation in Least Squares Regression.

Biometrika, 37 (1950).

3. Goldsmith, P.L. and Whitfield, H.W., Average Run Lengths in Cumulative Chart Quality Control Shemes. *Techmometrics 3*, pp. 11-20 (1961).
4. Harris, T. J. and Ross, W. H., Statistical Process Control Procedures for Correlated Observations. *Canadian Journal of Chemical Engineering 69*, pp. 48-57 (1991).
5. Johnson, R.A. and Bagshaw, M., The Effect of Serial Correlation on the Performance of CUSUM Tests. *Technometrics 16*, pp. 103-112 (1974).
6. Macgregor, J.F., A Different View of the Funnel Experiment. *Journal of Quality Technology 22*, pp. 255-259 (1990).
7. Vanbrackle, L.N., III and Reynolds, M.R., Jr., EWMA and CUSUM control charts in the presence of correlation (1994).
8. Yashchin, E., Performance of CUSUM Control Schemes for Serially Correlated Observations. *Technometrics*. *35*, pp. 37-52 (1993).