SENSITIVITY OF A CUSUM CHART FOR DISPERSION TO THE INDEPENDENCE ASSUMPTION

César A. Acosta-Mejía
Department of Industrial Engineering
Instituto Tecnológico Autónomo de México
México D.F. C.P. 01000

Joseph J. Pignatiello Jr.
Department of Industrial Engineering
Texas A&M University
College Station, TX 77843-3131, U.S.A.

ABSTRACT

Most control charts for variables data are constructed upon the assumption that the observations from a process are independent. In most applications however this assumption does not hold. The performance of process control schemes have to be evaluated in such circumstances. However most of the work published deal with shifts in the location of correlated processes. In this paper, the ARL performance of a CUSUM chart for process dispersion is analyzed, when the process does not comply with the assumption of independence. The ARL performance of the CUSUM chart for dispersion derived assuming independent observations is compared with that when the chart is used with autoregressive processes AR(1) with different degrees of correlation. It is found that the correlation structure increases the rate of false alarms of the CUSUM chart for σ . Also this chart is found much less sensitive in detecting changes in the process dispersion.

INTRODUCTION

Statistical Process Control (SPC) has been a widely used tool to monitor a process. It has proven to be very useful in monitoring many discrete processes. However it has failed to be as useful in the processes industries. In most of these industries observations show a high level of time dependency. This high level of dependency affects seriously the performance of commonly used control charts. Control charts may show a loss in sensitivity to detect shifts in the process parameters when observations are not independent.

The observations from many manufacturing processes are not independent. Observations sampled from a continuous flow process typically have a high degree of dependency over time. Processes found in chemical plants, waste-water plants, nuclear reactors, paper mills, and some food processing plants, generate sequences of dependent observations. Discrete part manufacturing processes where the frequency of sampling is very high can also produce correlated observations.

When control charts derived under the assumption of independent observations are applied to these types of processes, their performances are likely to be adversely affected. It is also probable that the loss in sensitivity to changes in the process is related to the magnitude of the dependency among observations.

To measure how sensitive a control chart is to correlated observations, ARL curves can be constructed. The ARL curves for both the independent and the correlated processes could then be compared. The difference between these curves is a measure of the effect of the correlation on the performance of the control chart.

Researchers have suggested different alternatives to monitor the location of correlated processes. Most agree that if the correlation among observations is inherent to the process it should be estimated and monitored. For monitoring the location of a correlated process a common approach is to monitor the residuals that result from fitting a model to the process. Harris and Ross (1991) suggest using an EWMA chart when observations are not independent. English (1997) presents a good review of different approaches for SPC when data is time dependent.

The performance of the CUSUM chart for monitoring location when observations are not independent has been studied by several authors. Johnson and Bagshaw (1974) and Goldsmith and Whitfield (1961) analyzed the difference in performance of the CUSUM chart for observations from an AR(1) process. Yaschin (1993) evaluated the run length performance of CUSUM charts when the observations are serially correlated. Their results agreed in concluding that positive (negative) lag 1 serial correlation increases (decreases) the false alarm rate. This conclusion indicates the necessity that the serial correlation structure must be taken into account when designing control charts.

Acosta (1995) presented a CUSUM chart for monitoring process dispersion called the Normal CP CUSUM chart for σ. The chart was derived under the assumption of normal and independent observations. In this work we evaluate the performance of this chart when the observations are serially correlated and follow a first-order autoregressive process. The average run length (ARL) for different shifts in the standard deviation are estimated by simulation and the resulting ARL curves are presented. By comparing the ARL curves when a assumption is violated to that when it is not, one can assess the effect of the violation on the performance of the control chart.

THE CP CUSUM CHART FOR DISPERSION

The CP CUSUM chart for dispersion proposed by Acosta (1995) can be expressed as follows. Let σ_0 be the in-control value for the process standard deviation and let $\sigma_a^- \neq \sigma_0$ and $\sigma_a^+ \neq \sigma_0$ be the process standard deviations that needs to be detected. If X_t represents the $t^{\underline{th}}$ observation, then the CP CUSUM chart for monitoring changes in the process standard deviation is given by:

$$\begin{aligned} SU_t &= \max\{0, Z_t - k^+ + SU_{t-1}\} && (1) \\ SL_t &= \max\{0, -Z_t + k^- + SL_{t-1}\} && (2) \end{aligned}$$

where

$$Z_t = \left[\frac{X_t - \mu}{\sigma_0}\right]^2 \qquad (3)$$

and the reference values, $k^\pm$, are defined as:

$$k^\pm = \frac{\log \delta^\pm}{(1 - (1/\delta^\pm))}, \qquad (4)$$

with $\delta^+ = \sigma_{a+}^2/\sigma_0^2$ and $\delta^- = \sigma_{a-}^2/\sigma_0^2$ representing a standardized increase and decrease, respectively, in the process variance that one wishes to detect quickly. The CP CUSUM for dispersion signals whenever SU_t or SL_t exceeds its respective control limit, h_u or h_l.

THE ARL PERFORMANCE OF AUTOCORRELATED PROCESSES

We have evaluated the magnitude of the effect of serial correlation on the performance of the normal CP CUSUM for dispersion. The ARL values were estimated by simulation. 10,000 runs were performed for each shift in the process standard deviation. For each run individual observations were sampled from a first-order autoregressive process (AR(1)).

A process is first-order autoregressive if, for $t \geq 2$, the observations can be represented by the model

$$X_t = \gamma + \phi X_{t-1} + \epsilon_t \qquad (5)$$

where $-1 < \phi < 1$ and ϵ_t are iid $\sim N(0, \sigma_e^2)$. For this process the mean and variance can be shown to be

$$\begin{aligned} E(X_t) &= \frac{\gamma}{1-\phi} \\ Var(X_t) &= \sigma_{X_t}^2 = \frac{\sigma_e^2}{1-\phi^2}. && (6) \end{aligned}$$

It can be seen that as the correlation coefficient ϕ increases (decreases) in absolute value the process variance increases (decreases). Also for any $\phi \neq 0$, $\sigma_{X_t}^2 > \sigma_e^2$. Thus, if the process is in-control, and ϕ is assumed to be zero, as ϕ increases in absolute value, the rate of false alarms should increase.

We evaluated the performance of the CP CUSUM chart for the situation where the chart used assumed an *iid* process but ϕ was different from zero. Different values of $\phi \neq 0$ are studied.

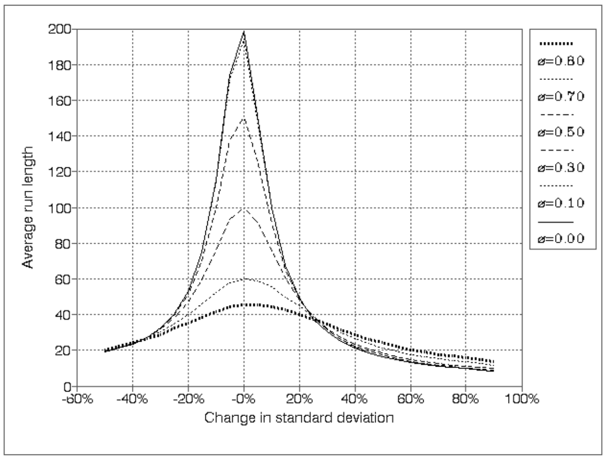

Figure 1: ARL curves for different $|\phi|$ values when the CP CUSUM chart is applied to an AR(1) process assuming ϕ to be known.

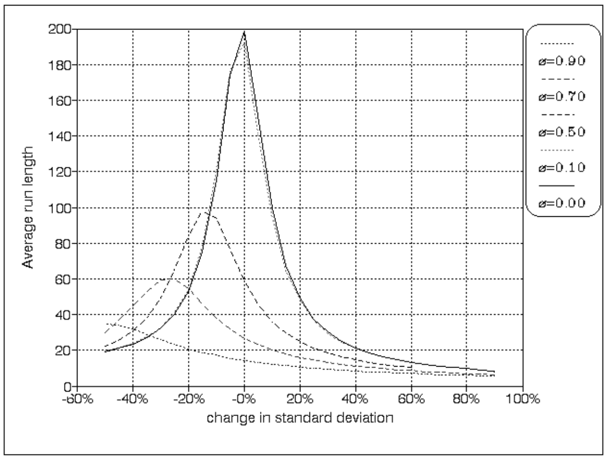

Figure 2: ARL curves for different $|\phi|$ values when the CP CUSUM chart is incorrectly applied to an AR(1) process.

Assume that the process is AR(1) with known ϕ, and apply the normal CP CUSUM chart derived for *iid* processes. In this case the chart used is defined by (1) and (2), but (3) is replaced by

$$Z_t = \left[\frac{X_t - \mu}{\sigma_{X_t}}\right]^2. \quad (7)$$

Assuming $\gamma = 0$ we estimate the performance of the normal CP CUSUM chart under these two circumstances by means of simulation. In all cases the chart parameters that were used were $h_u = 15.45$ and $h_l = 10.13$. These values give an on target ARL of 200 when the process is normally distributed with independent observations.

It is found that the correlation structure of the AR(1) process increases the rate of false alarms. This behavior is shown Figure 1. It can be seen that as the correlation increases the ARL curve becomes flatter. However for moderately to large changes in the standard deviation (i.e. greater than 40%) the ARL values found are not much different than those observed when the chart is used with $\phi = 0$.

In most applications however, the value of the correlation parameter ϕ is not known. In such case σ_{X_t} can not be used in (7) since it depends on the correct value of ϕ. Even if the user is aware that observations are not independent and the process is known to be AR(1), he or she might underestimate or overestimate the parameter ϕ and thus the process standard deviation σ_{X_t}.

Suppose that σ_{X_t} is underestimated in a magnitude such that $\sigma_0 = \sigma_e$ in equation (3). In this case the use of the CP CUSUM chart on an AR(1) process results in several ARL-biased curves. This is shown Figure 2. The bias in the ARL curve increases as the parameter ϕ increases in absolute value. In this situation the rate of false alarm increases, too. In addition the CUSUM chart becomes much less sensitive in detecting small to moderate decreases in the process dispersion than when it is used with an *iid* process. The chart's performance is greatly affected when $|\phi| > 0.50$.

However in both cases, when the CP CUSUM chart is applied on an AR(1) process and $|\phi| \leq 0.10$ the effect upon the chart's performance is very minor.

CONCLUSIONS

This paper shows the effect of serial correlation when a CUSUM chart derived under the assumption of independent observations is used to monitor the dispersion of an AR(1) process. When the process is recognized as an AR(1) and the correlation parameter ϕ is known, the overall effect is a decrease in the ARL values as ϕ increases in absolute value, resulting in an increase in the rate of false alarms. But when the process standard deviation is not known and is incorrectly estimated the CUSUM chart yields an ARL-biased curve. This bias in the ARL curve increases as the parameter ϕ increases in absolute value. In this case, the rate of false alarm increases as well. For $|\phi| \leq 0.10$ however, there is very little degradation in performance in both cases.

REFERENCES

1. Acosta (1995). "Monitoring Process Dispersion". Unpublished Ph.D. Thesis, Texas A&M University, College Station, Texas.
2. Bagshaw, M. and Johnson, R. A. (1975) "The effect of Serial Correlation on the Performance of CUSUM Tests II". *Technometrics* 17, pp. 73-80.
3. English J. R. (1997) "Monitoring Time Dependent Data". *Proceedings, $6^{\underline{th}}$ Industrial Engineering Research Conference* May 97, pp. 72-76.
4. Harris, T. S. and Ross W. H. (1991) "Statistical Process Control Procedures for Correlated Observations". *The Canadian Journal of Chemical Engineering* 69, pp. 48-57.
5. Yaschin, E. (1993). "Performance of CUSUM Control Schemes for Serially Correlated Observations". *Technometrics* 35, pp. 37-52.

Dr. Acosta-Mejía is professor in the Department of Industrial Engineering at the Instituto Tecnológico Autónomo de México, in Mexico City. He holds a Ph.D. in Industrial Engineering from Texas A&M University. His research interests include Statistical Process Control, Design of Experiments and Reliability Analysis. He is member of *ASQC, ASA and IIE.*

Dr. Pignatiello is an associate professor in the Department of Industrial Engineering at Texas A&M University. He has served as the Quality and Reliability Department Editor for *IIE Transactions* and currently is the *IIE Transactions* Department Editor for the Design of Experiments and Robust Design Department. He also serves on the editorial boards of *Quality Engineering*, the *Journal of Quality Technology* and the *International Journal of Industrial Engineering*.