Average Run Lengths of Special-Cause Control Charts for Autocorrelated Processes
자동상관인 공정에서 Special-Cause CUSUM 관리도의 ARL

최 성 운*
Sungwoon Choi[1)]

요 지

본 연구에서는 자동상관인 공정의 변화를 빠르게 탐지할 수 있는 Special-Cause CUSUM 관리도를 사용하여 다섯가지 시계열 모델에 대해 다음과 같은 연구를 수행한다. 첫째 ACF와 PACF로 파라미터에 따른 ARL의 변화를 쉽게 해석할 수 있는 방법과 둘째로 독립인 관측값에 적용하는 Hawkins(1992)의 ARL 간략계산법을 자동상관인 공정에서도 사용할 수 있는 기법을 제시하여 기존의 시뮬레이션을 이용한 ARL 계산법에 비해 빠르고도 정확한 값을 구한다. 끝으로 두가지 유형의 평균이동에 대한 ARL 변화를 각각 계산해 보아 그 효과를 비교분석한다.

1. Introduction

Traditional Statistical Process Control(SPC) assumes that consecutive observations from process are i.i.d.. Often in industrial practice, however, in continuous as well as discrete production processes, observations are actually autocorrelated. The effect of autocorrelation has been studied for several types of control charts. Johnson et al.(1974) and Yashchin(1993) have considered the effect of autocorrelation on CUSUM charts. Their studies show that as parameter value $\phi_1(\theta_1)$ increases in AR(1) model (MA(1) model), ARL decreases(increases). Maragas and Woodall(1992) have studied the effect of autocorrelation on the X-chart. It is shown that presence of positive(negative) parameter $\phi_1(\theta_1)$ in AR(1) model (MA(1) model) results in an increased number of false alarms from the control chart (i.e., type I error α increases) and negative(positive) parameter $\phi(\theta)$ in AR(1) model (MA(1) model(1) can result in unnecessarily wide control limits such that significant shifts in the process mean may go undetected (i.e., type II error β increases).

* Dept. of Industrial Engineering, KyungWon University

If autocorrelation is present and cannot be removed, then the usual interpretation of points outside standard control chart limits can be misleading. In this situation the approach taken by Alwan and Roberts(1988) is recommended. They introduced two charts, which they referred to as the common-cause control chart and the special-cause control chart. The common-cause chart as a plot of forecasted values that are determined by fitting the correlated process with an ARMA model, according to the procedures developed by Box et al.(1994). The special-cause chart is a traditional chart of residuals (i.e., the difference between the actual process values and their forecasts). English et al.(1991) suggested a similar approach using the residuals from Kalman filtering. This approach using state-space model allows a very general description of the stochastic and deterministic components of the noise processes as they appear as the system output (Goodwin et al.(1984), Goldsmith et al.(1971), Johnson et al.(1974), Harris et al.(1991)). This chart is used in traditional ways in detecting process deviations of any kind.

The ARL is the average number of observations required to obtain an observation outside of the control limits for a given shift in the mean. We normally desire the ARL to be large when no assignable cause has occurred and small when one has occurred. Some tables and graphs of the ARL are given in Lucas(1976). These provide the in-control and out-of control ARL of CUSUM charts with a range of reference values k=0.25 to 1.5, with the range of decision intervals h varying with k. While these tables are valuable for the ARL's corresponding to actual entries, their design is less than ideal in several respects. First, they cover a limited range of k and h values. Second, the spacings between the tabled values is quite large, which creates problems for evaluation the ARL for parameter values not explicitly listed. Other approaches involved substantial computations. Hawkins (1992) presents a relatively simple yet very accurate (typically within 3%) approximation for the ARL of a CUSUM chart, both when the process is in control and when it is out of control.

Harris and Ross(1991) investigated the effect of autocorrelation on the performance of a chart similar to the special-cause chart plots the CUSUM of residuals. They determined a simulated ARL for the CUSUM procedure when the process evolves according to an AR(1) model for various values the AR parameter, concluding that residual analysis is insensitive to shifts in the mean when the process is positively autocorrelated. Wardell et al.(1994) showed that the ARL and SRL of the special-cause control chart in ARMA(1,1) model are relatively smaller when the process is negatively rather than positively autocorrelated. In some instances they recommended the use of conventional control charts because they are at least as good, and often better in terms of ARL when the process is positively autocorrelated and are much easier to implement.

This paper suggests an alternative method which leads to easily computable approximations for ARL of special-cause CUSUM control charts for autocorrelated processes. The basic idea involves replacing a ARL approximations for i.i.d. sequence (Hawkins 1992) by same ones for the residuals of serially correlated sequence. In section 2, we introduce a Hawkins'(1992) approximation for ARL when the process is i.i.d.. In section 3, for the five time-series models we present new techniques to identify the impact of parameters on ARL using ACF and PACF, and a fast accurate approximation for ARL of special-cause CUSUM control charts, and two types of mean shifts as the assignable cause to be detected. Finally in section 4, we summarize and discuss possible research extensions.

2. ARL for iid Observations

Hawkins(1992) provides a computationally simple but accurate approximation for the ARL of CUSUM control chart. The starting point of the work was a closely-spaced table of ARL values. The table covered h values from 0 to 9 in steps of 0.5 and k values from -2 to 3 in steps of 0.125

The basic method for using this model as follows.

(i) For the selected h and k values obtain the necessary values of β_k, η_k, η_k^*, α_h, ζ_k, and ζ_h^* tables.

(ii) Using the values obtained in (i), compute $Y_{hk} = \alpha_h + \beta_k + \zeta_h \eta_k + \zeta_h^* \eta_k^*$.

(iii) Obtain ARL = $\frac{1}{\Phi(-y_{kh})}$ using a table or algorithm for the cumulative normal.

The ARL when the data depart from control is also of interest. In control, the data follow a N(μ, σ^2) distribution. Let us suppose that the mean and/or the standard deviation change. Specifically, suppose that the distribution of x_j changes to N($\mu+\Delta\sigma$, $\tau^2\sigma^2$). Writing the standardized quantity $U_j = \frac{x_j - \mu}{\sigma}$ as $U_j = \Delta + \tau W_j$, where $W_j \sim N(0,1)$, the cumulative sum of U_j can be expressed in terms of the W_j. The ARL of an $L^+(L^-)$ CUSUM chart using the U_j is the same as the in-control ARL of a CUSUM chart with reference value $\frac{k-\Delta}{\tau}\left(\frac{k+\Delta}{\tau}\right)$, decision interval $\frac{h}{\tau}$.

2.1 Mean Shifts

The in-control mean is 1,000 and the standard deviation is 50. For exactly normal data with no perturbation, ARL for a representative CUSUM scheme using k=0.250 and h=6 has an approximately 120.

Suppose that the standard deviation remains fixed and the mean changes by 12.5; 25; 50. The one-sided ARL's for detecting these upward shifts can be obtained.

Shift	12.5	25	50
Δ	0.250	0.500	1.000
Altered k	0	-0.25	-0.75
ARL	50.8	21.1	8.7
FIR ARL	38.2	13.4	6.0
Lucas' ARL	53	20	8.6

2.2 Mean Shifts With Group Means

A question often asked about CUSUM charts is whether one should plot individual observations or group means. Consider the problem discussed in 2.1 with group of size four. Evaluating the ARL for the three shifts of interest results in the following.

Shift	12.5	25	50
Δ	0.500	1.000	2.000
Altered k	-0.25	-0.75	-1.75
ARL	21.1	87	4.1
FIR ARL	12.4	6.0	2.4
Lucas' ARL	20	8.6	3.0

2.3 Simultaneous Shifts in Mean and Standard Deviation

We evaluate the response of the CUSUM chart to a simultaneous change in mean and standard deviation. Suppose that the process mean shifts to 1,040, and the standard deviation decreases to 35. In the general formulation these correspond to shifts of Δ=0.800 and τ=0.7. To get the ARL of the conventional CUSUM chart, we use reference value $\frac{0.250-0.800}{0.7}=-0.786$ and decision interval $\frac{6}{0.7}=8.57$. Entering these values into the program gives the ARL as 11.5 for a conventional CUSUM chart and 6.4 for a FIR CUSUM chart.

2.4 Standard Deviation Shifts

To investigate the out-of-control performance of a scale CUSUM chart, suppose that the standard deviation of the original x_j changes by a factor of ρ, while the mean remains fixed. Then,

$$U_j = \frac{x_j-\mu}{\sigma} \sim \rho N(0,1)$$

$$\sqrt{U_j} \sim \sqrt{\rho}\, N(0.822, (0.349)^2)$$

$$= N(0.822\sqrt{\rho}, (0.349)^2 \rho)$$

and

$$V_j = \frac{\sqrt{|U_j|}-0.822}{0.349} \sim \frac{N(0.822(\sqrt{\rho}-1), (0.349)^2\rho)}{0.349}$$

$$= N(2.355(\sqrt{\rho}-1), \rho)$$

The performance of the scale CUSUM chart can, therefore, be checked using the general formula with $\Delta = 2.355(\sqrt{\rho}-1)$ and $\tau = \sqrt{\rho}$.

Shift	100	150
ρ	1.41	1.73
τ	1.19	1.32
Δ	0.441	0.743
Altered k	-0.161	-0.373
Altered h	5.0	4.5
ARL	21.5	11.7
FIR ARL	14.7	7.4

Hawkins(1981) shows that on increase in σ leads to better performance than a decrease of the same magnitude. This is a highly desirable property in practice. Since one is normally concerned much more about an increases in σ than a decrease.

3. ARL for Serially-Correlated Observation

3.1 Impact of Parameters on ARL Using ACF and PACF

In a stationary autoregressive process of order p, a_t can be represented as a finite weighted sum of previous $\tilde{Z}_t$ as an infinite weighted sum

$$\tilde{Z}_t = \phi^{-1}(B)\, a_t$$

of previous a_t. Also an invertible moving average process of order q, $\tilde{Z}_t$ can e represented as a finite weighted sum of previous a_t as an infinite weighted sum

$$\theta^{-1}(B)\, \tilde{Z}_t = a_t$$

of previous $\tilde{Z}$.

The finite MA process has an ACF ρ_k is zero beyond a certain point, but since it is equivalent to an infinite AR process, its PACF ϕ_{kk} is infinite in extend and is dominated by damped exponentials and/or damped sine waves. Conversely, the AR process has a PACF ϕ_{kk} that is zero beyond a certain point, but its ACF ρ_k is infinite in extend and consists of a mixture of damped exponentials and/or damped sine waves. The spectrum of a MA process has an inverse relationship to the spectrum of the corresponding autoregressive process.

For the AR process, the ACF can be computed from $\tilde{Z}_t$. Because autocorrelation is a source of variability, in case the ACF is damped sine waves, serially-correlated observations are unstable, so ARL is smaller than the other case. Similarly, for the MA process, the PACF can be computed from a_t. In case the PACF is damped sine waves, noise components are random white, so ARL is greater than the other case.

As shown in Figure 1, the ACF decays exponentially to zero when ϕ_1 is positive, but decays exponentially to zero and oscillates in sign when ϕ_1 is negative. ARL of $\phi_1<0$ is smaller than the one of $\phi_1>0$. Contrary to its ACF, which cuts off ofter lag 1, the PACF of a MA(1) model tails off exponentially in one of two forms depending on the sign of θ_1. If alternating in sign, it begins with a positive value; otherwise, it decays on the negative side, as shown in Figure 2. ARL of $\theta_1<0$ is greater than the one of $\theta_1>0$. As shown in Figure 3, in region 1, the ACF remains positive and, in region 2, alternates in sign as it damps out. The ACF is a damps sine wave in region 3 and 4, the phase angle being less then 90° in region 4 and lying between 90° and 180° in region 3. ARL is small in order of region 2, 4, 3. As shown in Figure 4, in region 2, 3, 4, the PACF represents a damped sine wave, so ARL is great in order of region 2, 3, 4. As shown in Figure 5, exponential decay(ACF) is smooth if ϕ_1 is positive and alternates in ϕ_1 is negative. Furthermore, the sign of ρ_1 is determined by the sign of $(\phi_1 - \theta_1)$ and dictates from which side of zero the exponential decay(ACF) take place. When θ_1 is positive it is dominated by a smoothly damped exponential(PACF) which decays from a value of ρ_1, with sign determined by the sign $(\phi_1 - \theta_1)$. Similarly, when θ is negative, it is dominated by an exponential which oscillates as it decays from a value of ρ_1, with sign determined by the sign of $(\phi_1 - \theta_1)$. ARL is the smallest in region 3.

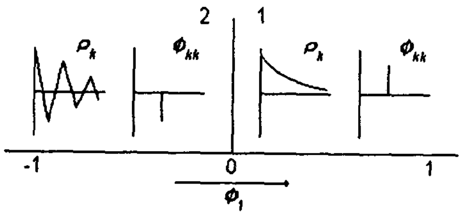

Figure 1
ACF ρ_k and PACF ϕ_{kk} for AR(1) model

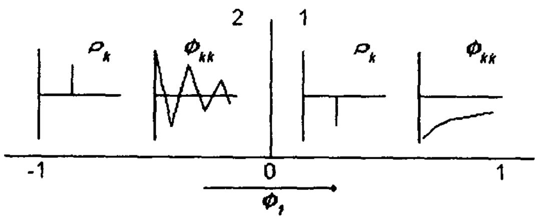

Figure 2
ACF ρ_k and PACF ϕ_{kk} for MA(1) Model

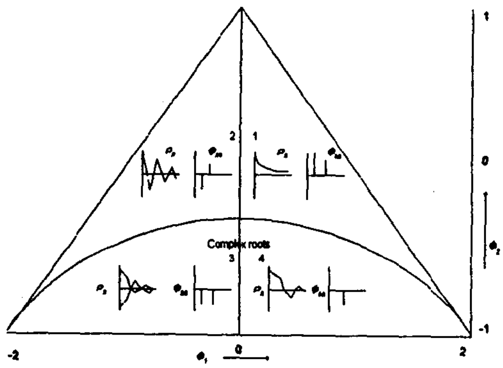

Figure 3
ACF ρ_k and PACF ϕ_{kk} for AR(2) model

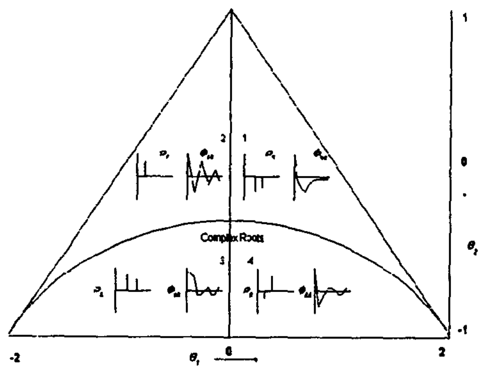

Figure 4
ACF ρ_k and PACF ϕ_{kk} for MA(2) model

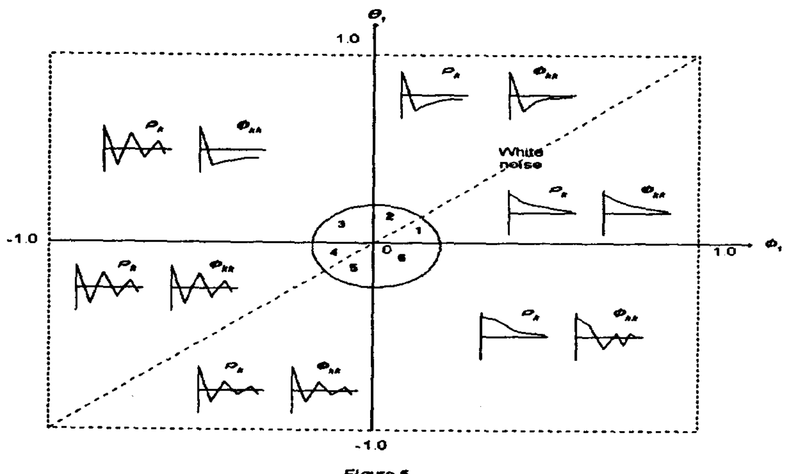

Figure 5
ACF ρ_k and PACF ϕ_{kk} for ARMA(1,1) model

3.2 ARL of Special-Cause CUSUM Control Charts

To compute the ARL, we consider a selected set of numerical results for typical time series model assuming that the random error terms, a_t, are normally distributed with mean 0 and variance σ_a^2. We choose five time series model to shown the effects of both the AR and MA parameters. The standard deviation of the observation is dependent on the model parameters and the standard deviation of the error terms as follows;

AR(1) $\sigma_Z = \sigma_a\sqrt{\dfrac{1}{1-\phi_1^2}}$

MA(1) $\sigma_Z = \sigma_a\sqrt{1+\theta_1^2}$

AR(2) $\sigma_Z = \sigma_a\sqrt{\left(\dfrac{1-\phi_2}{1+\phi_2}\right)\dfrac{1}{(1+\phi_2)^2-\phi_1^2}}$

MA(2) $\sigma_Z = \sigma_a\sqrt{1+\theta_1^2+\theta_2^2}$

ARMA(1,1) $\sigma_Z = \sigma_a\sqrt{\dfrac{1-2\phi_1\theta_1+\theta_1^2}{1-\phi_1^2}}$

where σ_Z is the standard deviation of the observations

Suppose $a_t \sim N(0, 1^2)$ and ARL = 120, k = 0.250, h = 6. we can use the same methodologies like 2.4 to compute ARL.

	AR(1) : ϕ_1 = -0.950	ARMA(1,1) : ϕ_1 = -0.950, θ_1 = 0.450
ρ	1.79	2.14
τ	1.34	1.46
Δ	0.796	1.090
Altered k	-0.407	-0.575
Altered h	4.5	4.0
ARL	10.9	7.5
FIR	7.0	4.3

3.3 Two Types of Process Mean Shifts

The shifts are measured in terms of the standard deviation of the observations σ_Z and the error terms σ_a.

In case the shifts are measured in terms of σ_Z (i) obtain altered k and h by using the same methodologies like 2.4 and (ii) compute ARL approximations according to σ_Z shifts like 2.1.

Consider the AR(1) : ϕ_1 = -0.950 model discussed in 3.2.

AR(1) : ϕ = -0.950, $a_t \sim N(0, 1^2)$

(i) altered k : -0.407 altered h : 4.5

(ii)

σ_Z shift	3.20	4.80
Δ	1.0	1.5
k	-1.407	-1.907
ARL	3.9	3.0
FIR ARL	2.3	1.8

In case the shifts are measured in terms of σ_a (i) obtain altered k according to σ_a shifts by using the same methodologies like 2.1 and (ii) compute ARL approximations like 2.4.

AR(1) : ϕ_1 = -0.950, $a_t \sim N(0, 1^2)$

(i) σ_a shift : 1.0 1.5

Δ 1.0 1.5

altered k : -0.75 -1.25

(ii)

Δ	-1.154	-1.527
k	4.5	4.5
ARL	4.7	3.6
FIR ARL	2.8	2.2

When to use the shifts of the standard deviations of the observations σ_Z, rather than that of the standard deviations of the error terms σ_a, ARL is smaller. It can be extended to investigate other types of shifts, including ramping shifts (i.e., gradual shifts in which the mean slowly increases and decreases to its new value), square impulses (in which the mean steps up to a new value for a time but then moves back down), and cyclical variations.

4. Conclusion

In this paper, we develop an efficient method which leas to easily computable approximations for ARL of special-cause CUSUM control chart for autocorrelated processes. These results show that for the AR(1) model the ARL is relatively smaller when $\phi_1<0$ rather than $\phi_1>0$ and for the MA(1) model the ARL is greater when $\theta_1<0$ rather than $\theta_1>0$. These facts can be also identified from the shapes of autocorrelation function and partial autocorrelation function. When to use the shifts of the standard deviation of the observations rather than that of the standard deviations of error terms, ARL is smaller.

This paper can be applied to monitor engineering process control problem to consider a systematic feedforward-feedback corrective action that will reduce the process deviation from the target.

References

Alwan, L.C. and Roberts, H.V.(1988), *"Time Series Modeling for Statistical Process Control,"* Journal of Business & Economic Statistics, 6, 87-95.

Box. G.E.P., Jenkins, G.H., and Reinsel, G.C.(1994), *Time Series Analysis: Forecasting and Control(3rd ed.),* Prentice-Hall.

English, J.R., Krishnamurthi, M., and Sastri, T.(1991), *"Quality Monitoring of Continuous Flow Processes,"* Computers and Industrial Engineering , 20, 251-260.

Goldsmith, P.C. and Whitfield, H.(1971), *"Average Run Lengths in Cumulative Chart Quality Control Schemes,"* Technometrics, 3, 11-20.

Goodwin, G.C., and Sin, K.S.(1984), *Adaptive Filtering Prediction and Control,* Prentice-Hall.

Harris, T.J., and Ross, W.H.(1991), *"Statistical Process Control Procedures for Correlated Observations,"* Canadian Journal of Chemical Engineering, 69, 48-57.

Hawkins, D.M.(1981), *"A CUSUM for a Scale Parameter,"* Journal of Quality Technology, 13, 228-231.

Hawkins, D.M.(1992), *"A Fast Accurate Approximation for Average Run Lengths of CUSUM Control Charts,"* Journal of Quality Technology, 24, 37-43.

Johnson, R.A. and Bagshaw, M.(1974), *"The Effect of Serial Correlation on the Performance of CUSUM Tests,"* Technometrics, 16, 103-112.

Lucas, J.M.(1967), *"The Design and Use of V-Mask Control Schemes,"* Journal of Quality Technology, 8, 1-11.

Maragah, H.D., and Woodall, W.H.(1992), *"The Effect of Autocorrelation on the Restrospective X-Chart."* Journal of Statistical Compution and Simulation, 46, 29-42.

Wardell D.G., Moskowitz, H., and Plante, R.D.(1994) *"Run-Length Distributions of Special-Case Control Charts for Correlated Processes,"* Technometrics, 36, 3-27.

Yashchin, E.(1993), *"Performance of CUSUM Control Schemes for Serially Correlated Observations,"* Technometrics, 36, 37-52