A Distribution-Free Tabular CUSUM Chart for Autocorrelated Data

SEONG-HEE KIM, CHRISTOS ALEXOPOULOS, and KWOK-LEUNG TSUI

School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332

JAMES R. WILSON

Department of Industrial Engineering, North Carolina State University Raleigh, NC 27695-7906

A distribution-free tabular CUSUM chart is designed to detect shifts in the mean of an autocorrelated process. The chart's average run length (ARL) is approximated by generalizing Siegmund's ARL approximation for the conventional tabular CUSUM chart based on independent and identically distributed normal observations. Control limits for the new chart are computed from the generalized ARL approximation. Also discussed are the choice of reference value and the use of batch means to handle highly correlated processes. The new chart is compared with other distribution-free procedures using stationary test processes with both normal and nonnormal marginals.

Key Words: Statistical Process Control; Tabular CUSUM Chart; Autocorrelated Data; Average Run Length; Distribution-Free Statistical Methods.

Biographical Note

Seong-Hee Kim is an Assistant Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. She is a member of INFORMS and IIE. Her e-mail and web addresses are `<skim@isye.gatech.edu>` and `<www.isye.gatech.edu/~skim/>`, respectively.

Christose Alexopoulos is an Associate Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He is a member of INFORMS. His e-mail address is `<christos@isye.gatech.edu>` and his web page is `<www.isye.gatech.edu/~christos/>`.

Kwok-Leung Tsui is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He is a regular member of ASQ. His e-mail address is <ktsui@isye.isye.gatech.edu>, and his web page is <www.isye.gatech.edu/people/faculty/Kwok_Tsui/>.

James R. Wilson is a Professor in the Department of Industrial Engineering at the North Carolina State University. He has served as head of the department since 1999. He is a member of INFORMS and IIE. His e-mail address is <jwilson@ncsu.edu>, and his web page is <http://www.ie.ncsu.edu/jwilson/>.

Introduction

Given a stochastic process to be monitored, a statistical process control (SPC) chart is used to detect any practically significant shift from the in-control status for that process, where the in-control status is defined as maintaining a specified target value for a given parameter of the monitored process—for example, the mean, the variance, or a quantile of the marginal distribution of the process. An SPC chart is designed to yield a specified value ARL_0 for the in-control average run length (ARL) of the chart—that is, the expected number of observations sampled from the in-control process before an out-of-control alarm is (incorrectly) raised. Given several alternative SPC charts whose control limits are determined in this way, one would prefer the chart with the smallest out-of-control average run length ARL_1, a performance measure analogous to ARL_0 for the situation in which the monitored process is in a specific out-of-control condition. If the monitored process consists of independent and identically distributed (i.i.d.) normal random variables, then control limits can be determined analytically for some charts such as the Shewhart and tabular CUSUM charts as detailed in Montgomery (2001).

It is more difficult to determine control limits for an SPC chart that is applied to an autocorrelated process; and much of the recent work on this problem has been focused on developing distribution-based (or model-based) SPC charts, which require one of the following:

1. The in-control and out-of-control versions of the monitored process must follow specific probability distributions.

2. Certain characteristics of the monitored process—such as such as the first- and second-order moments, including the entire autocovariance function—must be known.

Moreover, the control limits for many distribution-based charts can only be determined by trial-and-error experimentation. Of course, if the underlying assumptions about the probability distributions describing the target process are violated, then these charts will not perform as advertised. Another limitation is that determining the control limits by trial-and-error experimentation can be very inconvenient in practical applications—especially in circumstances that require rapid calibration of the chart and do not allow extensive preliminary experimentation on training data sets to estimate ARL_0 for various trial values of the control limits and other parameters of the chart. We illustrate these limitations of distribution-based charts in more detail in the next section, using an example from network intrusion detection.

The limitations of distribution-based procedures can be overcome by distribution-free SPC charts. Runger and Willemain (R&W) (1995) organize the sequence of observations of the monitored process into adjacent nonoverlapping batches of equal size; and their SPC procedure is applied to the corresponding sequence of batch means. They choose a batch size large enough to ensure that the batch means are approximately i.i.d. normal, and then they apply to the batch means one of the classical SPC charts developed for i.i.d. normal data. On the other hand, Johnson and Bagshaw (J&B) (1974) and Kim et al. (2005) present CUSUM-based methods that use raw observations instead of batch means. Computing the control limits for the latter two procedures requires an estimate of the variance parameter of the monitored process—that is, the sum of covariances at all lags. Nevertheless, these

CUSUM-based charts are distribution free since we can estimate the variance parameter using a variety of distribution-free techniques that are popular in the simulation literature; see Alexopoulos, Goldsman, and Serfozo (2005).

For first-order autoregressive processes, Kim et al. (2005) show that (i) their New CUSUM chart performs uniformly better than the J&B chart in terms of ARL_1 for a given target value of ARL_0; and (ii) the New CUSUM chart works better than the R&W Shewhart chart for small shifts. On the other hand, Kim et al. (2005) find that the R&W Shewhart chart performs better than the New CUSUM chart for large shifts. This is not surprising, given that a Shewhart-type chart is generally more effective than a CUSUM-type chart in detecting large shifts in processes consisting of independent normal observations. However, the R&W Shewhart chart may delay legitimate out-of-control alarms for processes with a pronounced correlation structure or large shifts; and in practice it is often difficult to determine a good choice for the batch size in the R&W Shewhart chart.

In this paper we formulate a distribution-free tabular CUSUM chart for monitoring an autocorrelated process. This new chart is a generalization of the conventional tabular CUSUM chart that is designed for i.i.d. normal random variables. Moreover to improve upon the performance of the J&B chart, our distribution-free tabular CUSUM chart incorporates a nonzero reference value into the monitoring statistic. For a reflected Brownian motion process with drift, Bagshaw and Johnson (1975) derive the expected first-passage time to a positive threshold; and they mention that this result can be used to approximate the ARL of a CUSUM chart with nonzero reference value. Combining this approximation with a generalization of the Brownian-motion approximation of Siegmund (1985) for the ARL of a CUSUM-based procedure that requires i.i.d. normal random variables, we design a distribution-free tabular CUSUM chart that can be used with raw correlated data or with batch means based on any batch size.

The rest of this article is organized as follows. The second section contains relevant background information, including a motivating example, notation, and assumptions. The third section presents the proposed distribution-free tabular CUSUM chart for autocorrelated processes. The fourth section contains an experimental comparison of the performance of the new procedure with that of existing distribution-free procedures based on the following test processes whose probabilistic behavior is typical of many practical applications of SPC procedures to autocorrelated processes:

1. the first-order autoregressive (AR(1)) process with lag-one correlation levels 0.0, 0.25, 0.5, 0.7, 0.9, 0.95, and 0.99; and

2. the sequence of waiting times spent in the queue for an $M/M/1$ queueing system with traffic intensities of 30% and 60% so that in steady-state operation, each configuration of the system has the following properties:

   a. the autocorrelation function of the process decays at an approximately geometric rate; and

   b. the marginal distribution of the process is markedly nonnormal, with an atom at zero and an exponential tail.

The final section summarizes the main findings of this work.

Background

In this section we give a motivating example from the area of intrusion detection in information systems to illustrate the emerging need for distribution-free SPC methods. Then we define notation and assumptions on the monitoring process for this article.

Motivating Example

The MIT Lincoln Laboratory simulated the environment of a real computer network to provide a test-bed of data sets for comprehensive evaluation of the performance of various intrusion detection systems. Ye, Li, Chen, Emran, and Xu (2001), Ye, Vilbert, and Chen (2003), and Park (2005) derive event-intensity (arrival-rate) data from log files generated by the Basic Security Module (BSM) of a Sun SPARC 10 workstation running the Solaris operating system and functioning as one of the components of the network simulated by the MIT Lincoln Laboratory. These authors consider a Denial-of-Service (DoS) attack on the Sun workstation that leaves trails in the audit data—in particular, they capture the activities on the machine through a continuous stream of audit events whose occurrence times are recorded in the log files.

Figure 1 shows event-intensity data (that is, the number of events in successive one-second time intervals) derived from the BSM audit file for an observation period of 12,000 seconds on a specific day in the data sets from the MIT Lincoln Lab. This data set is believed to be intrusion free. Since the Sun system performs a specific routine for creating a log file every 60 seconds, the graph in Figure 1 shows a repeated pattern every 60 seconds. After a careful analysis, Park (2005) separates the graph in Figure 1 into the cyclic and noise parts as shown in Figure 2.

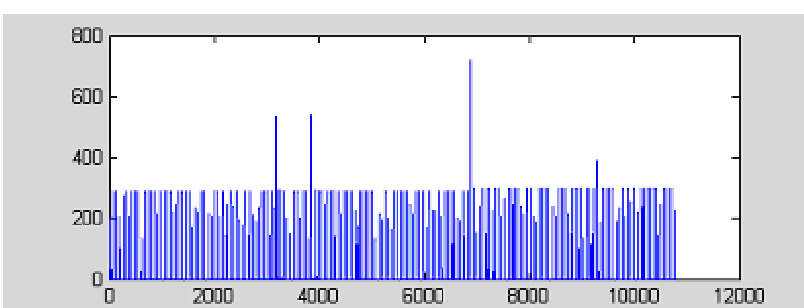

FIGURE 1. Example of Event Intensity from a BSM Audit File.

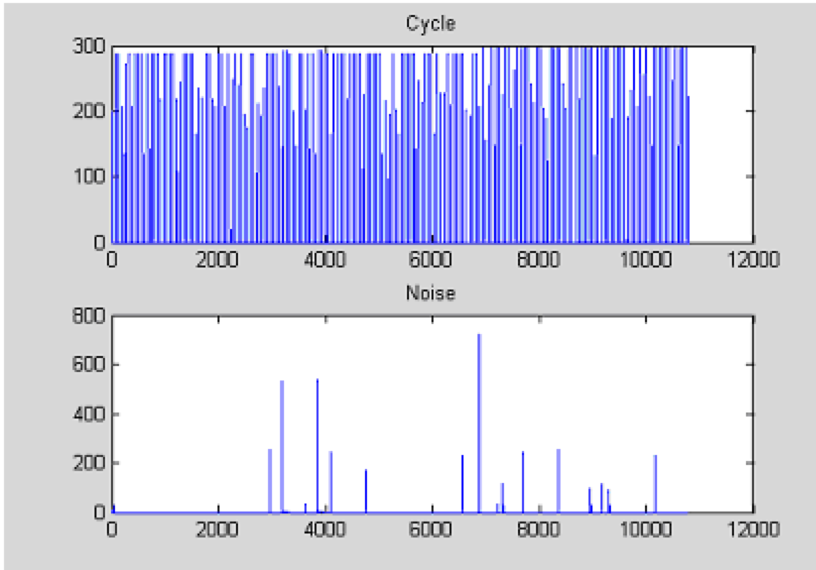

FIGURE 2. Example of Separated Event Intensity from a BSM Audit File.

For the detection of a DoS attack, the noise events must be monitored. One can observe that the noise data are very sparse—in particular, only 60 of the 12,000 one-second time intervals contained noise events not related to the generation of a log file so that the estimated probability of occurrence of at least one noise event in a given one-second time interval is only 0.005. No simple probability distributions (in particular, the Poisson and normal distributions) provided an adequate fit to the observed noise data because of its high standard deviation. For the sample of 60 noise-event counts associated with one-second time intervals containing at least one noise event as depicted in Figure 2, the sample mean is 81 and the sample standard deviation is 154, which is almost twice as large as the mean. Such anomalous behavior in the noise data strongly suggests that this process cannot be adequately represented by conventional univariate probability distributions; and ultimately Park fitted a Bézier distribution (Wagner and Wilson (1996)) to the nonzero noise-event counts displayed in the lower half of Figure 2 to drive a simulation-based performance evaluation of various intrusion detection procedures. For this application, it is clear that the currently used distribution-based SPC charts are inappropriate for detecting a DoS attack.

Notation and Assumptions

Suppose the discrete-time stochastic process $\{Y_i : i = 1, 2, \ldots\}$ to be monitored has a steady-state distribution with marginal mean $\mathrm{E}[Y_i] = \mu$ and marginal variance $\mathrm{Var}[Y_i] = \sigma_Y^2$. Specifically, we let μ_0 denote the in-control marginal mean. We let $\bar{Y}(n)$ denote the sample mean of the first n observations. The standardized CUSUM, $\mathcal{C}_n(t)$, is defined as

$$\mathcal{C}_n(t) \equiv \frac{\sum_{j=1}^{\lfloor nt \rfloor} Y_j - nt\mu}{\Omega_Y \sqrt{n}} \quad \text{for} \quad t \in [0, 1], \tag{1}$$

where: (i) $\lfloor \cdot \rfloor$ is the "floor" (greatest integer) function so that $\lfloor z \rfloor$ denotes the largest integer not exceeding z; and (ii) Ω_Y^2 is the variance parameter for the process $\{Y_i\}$, defined as

$$\Omega_Y^2 \equiv \lim_{n \to \infty} n \,\mathrm{Var}[\bar{Y}(n)] = \sum_{\ell=-\infty}^{\infty} \mathrm{Cov}(Y_i, Y_{i+\ell}),$$

and we assume that $0 < \Omega_Y^2 < \infty$. Let $\mathcal{W}(\cdot)$ denote a standard Brownian motion process on $[0, \infty)$ so that $\mathcal{W}(t)$ is normally distributed with $\mathrm{E}[\mathcal{W}(t)] = 0$ and $\mathrm{Cov}[\mathcal{W}(s), \mathcal{W}(t)] = \min\{s, t\}$ for $s, t \in [0, \infty)$.

For each positive integer n, the random function $\mathcal{C}_n(\cdot)$ is an element of the Skorohod space $D[0, 1]$, i.e., the space of functions on $[0, 1]$ that are right-continuous and have left-hand limits (Chapter 3 of Billingsley 1968). Our main assumption is that $\{Y_i : i = 1, 2, \ldots\}$ satisfies a Functional Central Limit Theorem (FCLT) (see Billingsley 1968, Chapter 4).

Assumption 1 *(FCLT) There exist finite real constants μ and $\Omega_Y^2 > 0$ such that as $n \to \infty$, the sequence of random functions $\{\mathcal{C}_n(\cdot) : n = 1, 2, \ldots\}$ converges in distribution to standard Brownian motion $\mathcal{W}(\cdot)$ in the Skorohod space $D[0, 1]$. Formally, we write*

$$\mathcal{C}_n(\cdot) \underset{n \to \infty}{\overset{\mathrm{D}}{\longrightarrow}} \mathcal{W}(\cdot),$$

where $\underset{n \to \infty}{\overset{\mathrm{D}}{\longrightarrow}}$ denotes convergence in distribution as $n \to \infty$.

Further, we assume that for every $t \in [0,1]$, the family of random variables $\{\mathcal{C}_n^2(t) : n = 1, 2, \ldots\}$ is uniformly integrable (see Billingsley 1968, Chapter 5).

Let

$$\mathcal{B}(t) = d_Y t + \Omega_Y \mathcal{W}(t) \quad \text{for} \quad t \in [0, \infty) \tag{2}$$

so that $\mathcal{B}(\cdot)$ denotes Brownian motion on $[0, \infty)$ with drift parameter d_Y and variance parameter Ω_Y^2 so that $\mathrm{E}[\mathcal{B}(t)] = d_Y t$ and $\mathrm{Var}[\mathcal{B}(t)] = \Omega_Y^2 t$ for all $t \geq 0$.

Tabular CUSUM for I.i.d. Normal Data

Given a monitored process consisting of i.i.d. normal random variables with marginal variance σ_Y^2, we see that the two-sided tabular CUSUM chart with reference value $K = k\sigma_Y$ and control limit $H = h\sigma_Y$ is defined by

$$S^{\pm}(n) = \begin{cases} 0, & \text{if } n = 0, \\ \max\{0,\ S^{\pm}(n-1) \pm (Y_n - \mu_0) - K\}, & \text{if } n = 1, 2, \ldots\,. \end{cases} \tag{3}$$

The interpretation of the $\pm$ notation in (3) is that (i) we have the initial values $S^+(0) = 0$, $S^-(0) = 0$; and (ii) for $n = 1, 2, \ldots$, we have $S^+(n) = \max\{0,\ S^+(n-1) + (Y_n - \mu_0) - K\}$ and $S^-(n) = \max\{0,\ S^-(n-1) - (Y_n - \mu_0) - K\}$. (Similar use of the $\pm$ notation is made throughout this article.) An out-of-control alarm is raised when the nth observation is taken if $S^+(n) \geq H$ or $S^-(n) \geq H$.

It is well known that the tabular CUSUM chart for i.i.d. normal data has nearly optimal sensitivity to a shift of magnitude $2K$; see p. 415 of Montgomery (2001). Therefore, if K (or k) is very small, then the chart is effective in detecting relatively small shifts but is less effective in detecting more meaningful shifts than a similar chart with a somewhat larger reference value. Table 1 shows ARLs of the tabular CUSUM chart with the reference parameter values $k = 0$ and $k = 0.5$. As expected, the tabular CUSUM chart with $k = 0$ is more effective in detecting shifts of size $0.25\sigma_Y$, but the chart with $k = 0.5$ detects any shift exceeding $0.25\sigma_Y$ much faster.

TABLE 1. ARLs of the Tabular CUSUM Chart When the Output Data Are I.i.d. Normal with Marginal Variance $\sigma_Y^2 = 1$, Where All Estimated ARLs Are Based on 1,000,000 Experiments.

Shift in Mean	Tabular CUSUM	
(Multiple of σ_Y)	$k = 0,\ h = 26.05$	$k = 0.5,\ h = 4.77$
0.00	370.08	368.76
0.25	100.88	121.20
0.50	52.47	35.22
0.75	35.46	16.18
1.00	26.80	9.92
1.50	18.04	5.51
2.00	13.64	3.86
2.50	11.00	3.00
3.00	9.24	2.48
4.00	7.04	1.96

Similarly, for autocorrelated data, we can expect that introducing a nonzero reference value into a CUSUM-type chart should improve the performance of the chart. The monitoring statistic of the J&B chart is the same as that of the tabular CUSUM chart but with reference value $K = 0$. Therefore, in the design of the distribution-free tabular CUSUM procedure, we incorporate the nonzero reference value K (or, equivalently, the reference parameter value k) into the monitoring statistic of the J&B chart. In the next section, we present the distribution-free tabular CUSUM chart with a nonzero reference value K and show how to determine the control limits.

A Distribution-Free Tabular CUSUM Procedure

For the one-sided monitoring statistics $S^+(n)$ and $S^-(n)$ defined in (3), we have the corresponding times at which an alarm is raised,

$$T_Y^{\pm} = \min\{n : S^{\pm}(n) \geq H \text{ and } n = 1, 2, \ldots\}. \tag{4}$$

In the rest of this section we discuss the computation of the average run length $\mathrm{E}[T_Y^+]$ for the in-control condition $\mathrm{E}[Y_i] = \mu_0$. A similar approach will yield the same final result for $\mathrm{E}[T_Y^-]$. To compute $\mathrm{E}[T_Y^+]$, we consider the following monitoring statistic that is closely related to $S^+(n)$ but is defined in a slightly different way,

$$S(n) = \begin{cases} 0, & \text{if } n = 0, \\ S(n-1) + (Y_n - \mu_0) - K, & \text{if } n = 1, 2, \ldots . \end{cases} \tag{5}$$

It is easy to see that $S^+(n)$ is always equal to $S(n) - \min\{S(\ell) : \ell = 0, 1, \ldots, n\}$ for $n = 1, 2, \ldots$. Set $d_Y = (\mathrm{E}[Y_j] - \mu_0) - K$. If n is sufficiently large, then it follows from Assumption 1 and the continuous mapping theorem (Billingsley 1968) that

$$\begin{aligned} S^+(n) &= S(n) - \min\{S(\ell) : \ell = 0, 1, \ldots, n\} \\ &\overset{\mathrm{D}}{\approx} d_Y n + \Omega_Y \sqrt{n}\mathcal{C}_n(1) - \inf\{d_Y tn + \Omega_Y \sqrt{n}\mathcal{C}_n(t) : 0 \le t \le 1\} \\ &\overset{\mathrm{D}}{\approx} d_Y n + \Omega_Y \sqrt{n}\mathcal{W}(1) - \inf\{d_Y tn + \Omega_Y \sqrt{n}\mathcal{W}(t) : 0 \le t \le 1\} \\ &\overset{\mathrm{D}}{=} d_Y n + \Omega_Y \mathcal{W}(n) - \inf\{d_Y u + \Omega_Y \mathcal{W}(u) : 0 \le u \le n\} \\ &\overset{\mathrm{D}}{=} \mathcal{B}(n) - \inf\{\mathcal{B}(u) : 0 \le u \le n\}, \end{aligned} \tag{6}$$

where: $\overset{\mathrm{D}}{=}$ denotes exact equality in distribution; $\overset{\mathrm{D}}{\approx}$ denotes approximate (asymptotically exact) equality in distribution; and $\mathcal{B}(\cdot)$ denotes Brownian motion with drift as defined in (2).

Now the stochastic process defined by

$$\mathcal{Z}(t) = \mathcal{B}(t) - \inf\{\mathcal{B}(u) : 0 \le u \le t\} \quad \text{for} \quad t \in [0, \infty) \tag{7}$$

has first-passage time to the threshold $H > 0$ given by $T_{\mathcal{Z}} = \inf\{t : t \ge 0 \text{ and } \mathcal{Z}(t) \ge H\}$. It follows from (6) and an argument similar to the proof of Proposition 2 of Kim, Nelson, and Wilson (2005) that if n is sufficiently large, then

$$\mathrm{E}[T_Y^+] \approx \mathrm{E}[T_{\mathcal{Z}}] = \begin{cases} H^2/\Omega_Y^2, & \text{if } d_Y = 0, \\ \dfrac{\Omega_Y^2}{2d_Y^2}\left[\exp\left(-\dfrac{2d_Y H}{\Omega_Y^2}\right) - 1 + \dfrac{2d_Y H}{\Omega_Y^2}\right], & \text{if } d_Y \ne 0, \end{cases} \tag{8}$$

where the formula for $\mathrm{E}[T_{\mathcal{Z}}]$ on the far right-hand side of (8) follows from Equation (2.1) of Bagshaw and Johnson (1975) or Theorem 3.1 of Darling and Siegert (1953).

For the situation in which the $\{Y_i\}$ are i.i.d. normal random variables, Siegmund (1985, p. 27) proposes an improvement to the approximation (8) for the expected first-passage time of the process $\{S^+(n) : n = 1, 2, \ldots\}$ to the control limit H. We formulate a distribution-free generalization of Siegmund's approximation to handle the case of observations that may be correlated or nonnormal as follows:

$$\mathrm{E}[T_Y^+] \approx \begin{cases} \dfrac{\Omega_Y^2}{2d_Y^2}\left\{\exp\left[-\dfrac{2d_Y(H+1.166\Omega_Y)}{\Omega_Y^2}\right] - 1 + \dfrac{2d_Y(H+1.166\Omega_Y)}{\Omega_Y^2}\right\}, & \text{if } d_Y \neq 0, \\ \left(\dfrac{H+1.166\Omega_Y}{\Omega_Y}\right)^2, & \text{if } d_Y = 0, \end{cases} \tag{9}$$

where the drift parameter $d_Y = (\mathrm{E}[Y_i] - \mu_0) - K$. If the monitored process is in control, then the right-hand side of (9) yields our approximation to $\mathrm{E}[T_Y^+]$ when we take $d_Y = -K = -k\sigma_Y$. Finally, considerations of symmetry in the definitions (3) of the one-sided process-monitoring statistics $S^+(n)$ and $S^-(n)$ and of their respective first-passage times defined by (4) reveal that $\mathrm{E}[T_Y^+] = \mathrm{E}[T_Y^-]$. To derive our new SPC chart, we determine the control limits based on (9) since the approximation is slightly more accurate than (8). It follows that the distribution-free tabular CUSUM has the following formal algorithmic statement:

Distribution-Free Tabular CUSUM Procedure

1. Choose K and a target two-sided ARL_0. Then, calculate H, the solution to the equation

$$\frac{\Omega_Y^2}{2K^2}\left\{\exp\left[\frac{2K(H+1.166\Omega_Y)}{\Omega_Y^2}\right] - 1 - \frac{2K(H+1.166\Omega_Y)}{\Omega_Y^2}\right\} = 2\mathrm{ARL}_0. \tag{10}$$

2. Raise an out-of-control alarm after the nth observation if $S^+(n) \geq H$ or $S^-(n) \geq H$.

Any simple search method such as the bisection method can be used to solve (10).

Determination of Parameters

The control limits of the distribution-free tabular CUSUM chart depend on the reference value K and the target value ARL_0 for the in-control average run length. Here, we search for the choice of K that guarantees good performance for the distribution-free tabular CUSUM chart by experiments, based on a stationary first-order autoregressive (AR(1)) process defined as follows:

$$Y_j = \mu + \varphi_Y(Y_{j-1} - \mu) + \varepsilon_j \quad \text{for} \quad j = 1, 2, \ldots, \tag{11}$$

where: (i) $\{\varepsilon_j : j = 1, 2, \ldots\} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, \sigma_\varepsilon^2)$; (ii) we take $-1 < \varphi_Y < 1$ to ensure that (11) defines a stationary AR(1) process; and (iii) we take $Y_0 \sim \mathrm{N}(\mu, \sigma_Y^2)$ to ensure that the process $\{Y_j\}$ starts in steady-state operation. If we take $\sigma_\varepsilon^2 = 1 - \varphi_Y^2$, then the marginal variance of the process $\{Y_j\}$ is $\sigma_Y^2 = 1$.

Naturally, one would consider using (i) $K = k\sigma_Y$, which is the choice for the tabular CUSUM chart with i.i.d. normal data; or (ii) $K = k\Omega_Y$, which seems to be the natural generalization of (i) for correlated data. The accuracy of (9) depends on whether $S^+(n)$ or $S^-(n)$ behaves like the process $\mathcal{Z}(\cdot)$ defined by (7). If K is too large, then $S^+(n)$ hits zero too frequently; and in this situation we have found that the convergence in distribution described by display (6) is too slow for the approximation (9) to yield acceptable accuracy. Therefore, K should not be too large, but at the same time it should not be too close to zero to ensure that the chart is sensitive to meaningful shifts. In practice, observations from the monitored process are likely to be positively correlated; and for positively autocorrelated data, the variance parameter Ω_Y^2 is often substantially larger than the marginal variance σ_Y^2. For example, the AR(1) process (11) with autoregressive parameter $\varphi_Y = 0.9$ and marginal variance $\sigma_Y^2 = 1$ has variance parameter $\Omega_Y^2 = \sigma_Y^2(1 + \varphi_Y)/(1 - \varphi_Y) = 19$. So, we set $K = k\sigma_Y$ rather than $K = k\Omega_Y$.

To find a good choice of the parameter k to yield an effective reference value $K = k\sigma_Y$, we set the two-sided ARL_0 equal to 10,000; or, equivalently, for both of the one-sided tests based on $S^{\pm}(n)$, we assigned the target in-control one-sided average run length equal to 20,000. Then we computed H from (10) with $k \in \{0.00001, 0.01, 0.03, 0.05, 0.1, 0.5\}$; and we recorded the experimentally observed two-sided ARL_0. As shown in Table 2 for AR(1) processes with $\varphi_Y \in \{0.25, 0.9\}$, the experimentally observed two-sided ARL_0's are close to the target two-sided $\mathrm{ARL}_0 = 10{,}000$ even for high correlation when k is small—say, $k \in \{0.00001, 0.01, 0.03\}$. But for large k (say, $k \geq 0.5$), the accuracy of the approximation falls off significantly even for a small correlation such as $\varphi_Y = 0.25$. We recommend the value $k = 0.1$ on account of the following considerations:

1. The reference parameter k should not be too large.

2. The reference parameter k should not be too close to zero.

3. In most practical applications of SPC charts, the lag-one correlation of the monitored process is rarely larger than 0.9.

The choice $k = 0.1$ ensures that the actual ARL_0 will be close to the target ARL_0 for small to medium correlation. However, the accuracy of the generalized approximation (9) to the ARL breaks down for high correlation, resulting in a conservative control limit H. As one can see in Table 2 for the case in which $\varphi_Y = 0.9$, the experimental two-sided ARL_0 is 13,468 when the target ARL_0 is 10,000. In next subsection, we present a method that finds a control limit H that ensures the actual ARL_0 is close to the target value.

Method for Handling Processes with High Correlation

The distribution-free tabular CUSUM chart incorporates a method for handling processes with excessively high correlation. In particular, there is substantial experimental evidence to show that when the distribution-free tabular CUSUM chart is applied to the original

TABLE 2. Experimental Two-Sided ARL_0 of the Distribution-Free Tabular CUSUM Chart with the Generalized Approximation (9) for an AR(1) Process, with Estimated ARLs Based on 5,000 Experiments.

k	$\varphi_Y = 0.25$	$\varphi_Y = 0.9$
0.00001	10255	11263
0.01	10578	11205
0.03	10399	11100
0.05	10245	11032
0.1	10264	13468
0.5	14841	34203

(unbatched) data $\{Y_i\}$, the procedure will only work as intended (that is, deliver an average run length approximate equal to ARL_0 for the in-control condition) when

$$\varphi_Y = \mathrm{Corr}(Y_i, Y_{i+1}) \le \zeta = 0.5; \tag{12}$$

see also Bagshaw and Johnson (1975). On the other hand, if the upper limit (12) on the lag-one correlation is not satisfied, then for an appropriate batch size $m > 1$, we compute batch means

$$\bar{Y}_i(m) = \frac{1}{m} \sum_{\ell=(i-1)m+1}^{im} Y_\ell \quad \text{for} \quad i = 1, 2, \ldots, b = \lfloor n/m \rfloor, \tag{13}$$

where we take m just large enough to ensure that the lag-one correlation between batch means will satisfy the requirement

$$\varphi_{\bar{Y}(m)} \equiv \mathrm{Corr}[\,\bar{Y}_i(m), \bar{Y}_{i+1}(m)] \le \zeta. \tag{14}$$

Then we apply the distribution-free tabular CUSUM procedure to the batch means process $\{\bar{Y}_j(m) : j = 1, \ldots, b\}$ with variance parameter given by

$$\Omega^2_{\bar{Y}(m)} = \Omega^2_Y / m. \tag{15}$$

The remainder of this section details the computation of the batch size m required to satisfy the upper limit (14) on the lag-one correlation of the batch means that will be used as the basic data items to which the distribution-free tabular CUSUM chart may properly be applied.

Suppose we are given a realization $\{Y_i : i = 1, \ldots, n\}$ of the original (unbatched) process from which we calculate the sample statistics

$$\left.\begin{aligned} \bar{Y}(n) = n^{-1}\sum_{i=1}^{n} Y_i, \quad S^2 = (n-1)^{-1}\sum_{i=1}^{n}[Y_i - \bar{Y}(n)]^2, \\ \widehat{\varphi}_Y = (n-1)^{-1}\sum_{i=1}^{n-1}[Y_i - \bar{Y}(n)][Y_{i+1} - \bar{Y}(n)]\Big/S^2. \end{aligned}\right\} \tag{16}$$

Under the assumption that $\{Y_i\}$ is a stationary AR(1) process with autoregressive parameter $\varphi_Y \in (-1, +1)$, we have

$$\sqrt{n}(\widehat{\varphi}_Y - \varphi_Y) \underset{n\to\infty}{\overset{\mathrm{D}}{\longrightarrow}} N(0, 1 - \varphi_Y^2); \tag{17}$$

see Theorem 8.2.1 and pp. 404–405 of Fuller (1996). Unfortunately, it is well known that the convergence to normality in (17) can be very slow when φ_Y is close to one. In particular, Figure 8.2.1 of Fuller (1996) clearly reveals the nonnormality of $\widehat{\varphi}_Y$ for the sample size $n = 100$ with $\varphi_Y = 0.9$.

Applying the delta method (Stuart and Ord 1994) to (17), we propose using the arc sine transformation of $\widehat{\varphi}_Y$,

$$\mathcal{S} = \sin^{-1}(\widehat{\varphi}_Y),$$

to test for the condition (12). From (17) and Corollary A.14.17 of Bickel and Doksum (1977), we obtain the asymptotic property

$$\sqrt{n}\left[\sin^{-1}(\widehat{\varphi}_Y) - \sin^{-1}(\varphi_Y)\right] \underset{n\to\infty}{\overset{\mathrm{D}}{\longrightarrow}} N(0, 1). \tag{18}$$

Thus when n is large, $\sin^{-1}(\,\widehat{\varphi}_Y\,)$ is approximately normal with mean $\sin^{-1}(\varphi_Y)$ and variance $1/n$.

We use the approximation

$$\sin^{-1}(\widehat{\varphi}_Y) \ \dot{\sim}\ \mathrm{N}\Big[\sin^{-1}(\varphi_Y), 1/n\Big]$$

to test the hypothesis (12) at the level of significance $\alpha_{\text{cor}} = 0.01$ by checking for the condition that the $100(1-\alpha_{\text{cor}})\%$ upper confidence limit for $\sin^{-1}(\varphi_Y)$ does not exceed the threshold $\sin^{-1}(\zeta)$. If we find

$$\sin^{-1}(\widehat{\varphi}_Y) + \frac{z_{1-\alpha_{\text{cor}}}}{\sqrt{n}} \le \sin^{-1}(\zeta) \iff \widehat{\varphi}_Y \le \sin\left[\sin^{-1}(\zeta) - \frac{z_{1-\alpha_{\text{cor}}}}{\sqrt{n}}\right] \tag{19}$$

(with $z_{1-\alpha_{\text{cor}}} = z_{0.99} = 2.33$), then we conclude that the original unbatched process $\{Y_i\}$ satisfies (12) and no batching is required before applying the distribution-free tabular CUSUM procedure.

If the condition (19) is not satisfied, then we compute the required batch size according to

$$m = \left\lceil \ln\left\{\sin\left[\sin^{-1}(\zeta) - \frac{z_{1-\alpha_{\text{cor}}}}{\sqrt{n}}\right]\right\}\Big/\ln(\widehat{\varphi}_Y)\right\rceil; \tag{20}$$

we compute the batch means (13) for batches of size m; and finally we apply the distribution-free Tabular CUSUM chart to the resulting batch means process. Note that in (20), $\lceil\cdot\rceil$ denotes the "ceiling" function so that $\lceil z\rceil$ is the smallest integer not less than z.

Remark. The basis for the batch size formula (20) is the approximation

$$\varphi_{\bar{Y}(m)} \approx \varphi_Y^m \tag{21}$$

as detailed in Appendix B of Steiger et al. (2005).

Experiments

In this section, we compare the performance of our procedure with those of other distribution-free SPC procedures designed for autocorrelated data. The comparisons are based on a stationary AR(1) model (11) and the queue waiting times observed in an $M/M/1$ queue.

We compare the distribution-free Tabular CUSUM with distribution-free SPC procedures due to Johnson and Bagshaw (1974), Kim et al. (2005), and Runger and Willemain (1995).

The J&B Two-Sided Chart: Define $S_n^{\pm} \equiv \max\{S_{n-1}^{\pm} \pm (Y_n - \mu_0), 0\}$ for $n \geq 1$, with $S_0^+ = 0$ and $S_0^- = 0$. Choose the target two-sided ARL_0, and set $H = \Omega_Y \sqrt{2\text{ARL}_0}$. Give an out-of-control signal after the nth observation if $S_n^+ > H$ or $S_n^- > H$.

The New CUSUM Chart: Choose a target ARL_0 and determine the control limit $H = \Omega_Y(\sqrt{\text{ARL}_0} - 1.166)$. Raise an out-of-control signal after the nth observation if

$$\left| \sum_{j=1}^{n} (Y_j - \mu_0) \right| \geq H.$$

The R&W Shewhart Chart: Find a batch size m such that the lag-one autocorrelation of the batch means is approximately 0.1. Choose a target ARL_0 and find z_{ON} such that

$$\frac{m}{1 - \Phi(z_{\text{ON}}) + \Phi(-z_{\text{ON}})} = \text{ARL}_0.$$

Then give an out-of-control signal if $|\text{UBM}_i| \geq z_{\text{ON}} \cdot \sigma_{\text{UBM}}$, where UBM_i is the ith uncorrelated batch mean (UBM) and σ^2_{UBM} is the marginal variance of the UBMs (which is assumed to be known).

It is important to recognize that all the experimental results reported in this article are based on the assumption that the variance parameter Ω_Y^2 is known. In many practical applications, this quantity must be estimated from a training data set; and it is unclear how the performance of the selected SPC procedures will be affected by estimation of the variance parameter. Nevertheless, the experimental results reported below provide some basis for ongoing research on the problem of developing SPC procedures for autocorrelated processes that are effective in practice.

AR(1) Processes

For the AR(1) process (11), the marginal variance is

$$\sigma_Y^2 = \frac{\sigma_\varepsilon^2}{1-\varphi_Y^2}; \tag{22}$$

the lag-ℓ covariance is

$$\mathrm{Cov}(Y_i, Y_{i+\ell}) = \sigma_Y^2 \varphi_Y^{|\ell|} = \frac{\sigma_\varepsilon^2 \varphi_Y^{|\ell|}}{1-\varphi_Y^2} \quad \text{for} \quad \ell = 0, \pm 1, \pm 2, \ldots\ ; \tag{23}$$

and the variance parameter is

$$\Omega_Y^2 = \sigma_Y^2 \left(\frac{1+\varphi_Y}{1-\varphi_Y} \right) = \frac{\sigma_\varepsilon^2}{(1-\varphi_Y)^2}. \tag{24}$$

In the experiments reported below, the marginal variance of Y_i is set to $\sigma_Y^2 = 1$; therefore, $\sigma_\varepsilon^2 = 1 - \varphi_Y^2$. The shift varies over 0, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3, and 4 in multiples of σ_Y, and the coefficient φ_Y is set to $0, 0.25, 0.5, 0.7, 0.9, 0.95$, and 0.99.

For R&W, we consider two different values for the batch size: (i) a batch size m_1 that yields a lag-one correlation between batch means of approximately 0.1; and (ii) a batch size m^* that minimizes the mean-squared error of the nonoverlapping batch means estimator of the variance parameter Ω_Y^2,

$$\widehat{\Omega}_Y^2 = \frac{m}{b-1} \sum_{i=1}^{b} [\bar{Y}_i(m) - \bar{Y}(n)]^2$$

(Chien, Goldsman, and Melamed 1997). The asymptotically optimal batch size m^* for the AR(1) process was derived by Carlstein (1986):

$$m^* = \left\{ \frac{2|\varphi_Y|}{1-\varphi_Y^2} \right\}^{2/3} n^{1/3}.$$

Since the target ARL_0 for a shift of zero is 10,000, we use $n = 10{,}000$ in the above equation to compute m^*. For all configurations of the AR(1) process, we perform 5,000 independent replications of each SPC procedure.

Table 3 displays the estimated ARLs for small to medium values of the lag-one correlation—that is, $\varphi_Y \in \{0, 0.25, 0.5\}$. In each row of the table, the boxed entry is the best (smallest) value of ARL_1 for the combination of the shift $(\mu - \mu_0)/\sigma_Y$ and the lag-one correlation φ_Y that defines the associated configuration of the test process $\{Y_i\}$. As expected, among the three CUSUM-type charts we consider in this article, the distribution-free tabular CUSUM always outperforms the other two charts. However, the R&W chart is more efficient than the distribution-free tabular CUSUM in detecting large shifts. This is not surprising since the R&W chart does not require a large batch size for AR(1) processes with small to medium lag-one correlation; and a Shewhart-type chart is usually more effective in detecting large shifts compared to a CUSUM-type chart.

Table 4 displays estimated ARLs for large values of the lag-one correlation—that is, $\varphi_Y \in \{0.7, 0.9, 0.95, 0.99\}$. For large lag-one correlation, we test the performance of the distribution-free tabular CUSUM with batching as well as without batching. Using the method described in the previous section, we chose the batch size for the distribution-free tabular CUSUM procedure so that the lag-one correlation of the batch means is approximately 0.5. As shown in Table 4, batching helps in the getting ARL_0 close to its target value of 10,000. There is some loss in the performance of ARL_1, but this is not that significant because the batch size is not that large.

For large correlation, the distribution-free tabular CUSUM does not always perform better than the other two CUSUM-type charts. There are a few cases in which New CuSum has a smaller value of ARL_1 than the distribution-free tabular CUSUM for small shifts. Both New CuSum and the distribution-free tabular CUSUM are more effective in detecting small shifts compared with R&W. However, for large shifts, R&W still performs better than the three CUSUM-type charts. However, when $\varphi_Y = 0.99$, R&W requires an excessive batch size; and then R&W requires one full batch even for large shifts. This delays legitimate out-of-control alarms and degrades the performance of the chart. This problem is demonstrated

TABLE 3. Two-Sided ARLs in Terms of Number of Raw Observations for an AR(1) Process with Small or Medium φ_Y and $\sigma_Y^2 = 1$.

φ_Y	Shift	J&B	New CuSum	Dist.-Fr. CUSUM	R&W	
					$m = 1$	
0	0	10112	10194	9585	9843	
	0.25	562	404	178	6390	
	0.5	284	202	72	2776	
	0.75	190	135	45	1164	
	1	142	102	33	520	
	1.5	95	68	21	119	
	2	71	52	16	34	
	2.5	57	42	13	12	
	3	48	35	11	5	
	4	36	27	8	2	
					$m_1 = 4$	$m^* = 15$
0.25	0	10182	10145	10846	9822	9985
	0.25	726	518	270	4345	1863
	0.5	366	261	111	1157	304
	0.75	244	174	69	366	81
	1	183	131	50	131	34
	1.5	123	87	32	28	16
	2	92	66	24	10	15
	2.5	74	53	19	6	15
	3	62	44	16	4	15
	4	46	33	12	4	15
					$m_1 = 8$	$m^* = 27$
0.5	0	10377	10086	11356	9985	9884
	0.25	973	697	434	4177	2062
	0.5	492	350	180	1162	382
	0.75	327	231	112	364	113
	1	247	174	82	138	51
	1.5	164	116	53	33	29
	2	123	86	39	14	27
	2.5	99	69	31	9	27
	3	82	57	26	8	27
	4	62	43	19	8	27

TABLE 4. Two-Sided ARLs in Terms of Number of Raw Observations for an AR(1) Process with High φ_Y and $\sigma_Y^2 = 1$.

φ_Y	Shift	J&B	New CuSum	Dist.-Fr. CUSUM		R&W	
					$m = 3$	$m_1 = 19$	$m^* = 43$
0.7	0	10452	10133	12252	11376	10087	9831
	0.25	1333	959	**718**	729	4105	2661
	0.5	674	478	**301**	310	1106	575
	0.75	453	319	187	198	372	**184**
	1	340	239	136	144	148	**87**
	1.5	227	159	87	94	**44**	47
	2	170	119	64	69	**24**	43
	2.5	136	95	51	55	**20**	43
	3	113	79	42	46	**19**	43
	4	85	60	31	34	**19**	43
					$m = 7$	$m_1 = 58$	$m^* = 97$
0.9	0	10957	10310	13256	11668	9949	9836
	0.25	2410	1761	1746	**1728**	4925	4126
	0.5	1243	880	**755**	**755**	1648	1225
	0.75	830	590	475	481	639	**470**
	1	623	438	342	352	304	**233**
	1.5	415	292	221	227	**109**	116
	2	311	219	162	167	**68**	98
	2.5	248	174	128	133	**59**	97
	3	208	145	106	111	**58**	97
	4	155	109	78	83	**58**	97
					$m = 15$	$m_1 = 118$	$m^* = 157$
0.95	0	11286	10772	13650	12032	9992	10172
	0.25	3404	**2556**	2880	2754	5516	4997
	0.5	1772	1269	1269	**1250**	2065	1790
	0.75	1190	850	796	792	892	**777**
	1	901	634	576	577	461	**411**
	1.5	595	421	370	377	**190**	199
	2	446	314	270	278	**131**	163
	2.5	357	251	214	223	**119**	157
	3	297	209	177	185	**118**	157
	4	223	156	131	139	**118**	157
					$m = 74$	$m^* = 463$	$m_1 = 596$
0.99	0	12911	12021	15552	12735	10031	10200
	0.25	7264	**5897**	7286	6735	7161	6849
	0.5	4031	**2978**	3678	3383	3717	3458
	0.75	2727	1983	2347	2240	2021	**1924**
	1	2047	1474	1727	1641	1256	**1218**
	1.5	1379	970	1105	1065	**647**	729
	2	1025	721	806	794	**501**	614
	2.5	815	577	637	636	**469**	598
	3	675	474	524	530	**463**	596
	4	504	**355**	387	402	463	596

more clearly in the following example involving waiting times in a single-server queueing system.

M/M/1 Queue Waiting Times

In an $M/M/1$ queueing system, we let A_i denote the interarrival time between the customers numbered $i-1$ and i (with $A_0 \equiv 0$) so that $\{A_i : i = 1, 2, \ldots\} \overset{\text{i.i.d.}}{\sim}$ Exponential(μ_A) and $\mathrm{E}[A_i] = \mu_A$; moreover, we let B_i denote the service time of the ith customer so that $\{B_i : i = 1, 2, \ldots\} \overset{\text{i.i.d.}}{\sim}$ Exponential(μ_B) and $\mathrm{E}[B_i] = \mu_B$. If Y_i denotes the waiting time in the queue for the ith customer in this single-server queueing system, then we see that

$$Y_{i+1} = \max\{0, Y_i + B_i - A_{i+1}\} \quad \text{for} \quad i = 1, 2, \ldots \ .$$

As detailed, for example, in Section 4.2 of Steiger and Wilson (2001), the $M/M/1$ queue waiting times $\{Y_i : i = 1, 2, \ldots\}$ constitute a test process with highly nonnormal marginals and an autocorrelation function that decays approximately at a geometric rate. In terms of the arrival rate $\lambda = 1/\mu_A$, the service rate $\nu = 1/\mu_B$, and traffic intensity $\tau = \lambda/\nu$, the process $\{Y_i\}$ has marginal distribution function

$$F_Y(y) \equiv \Pr\left\{Y_i \leq y\right\} = \begin{cases} 0, & y < 0, \\ 1 - \tau, & y = 0, \\ 1 - \tau e^{-(\nu-\lambda)y}, & y > 0, \end{cases} \tag{25}$$

so that the marginal mean and variance are given by

$$\mu = \mathrm{E}[Y_i] = \frac{\tau^2}{\lambda(1-\tau)} \quad \text{and} \quad \sigma_Y^2 = \mathrm{Var}[Y_i] = \frac{\tau^3(2-\tau)}{\lambda^2(1-\tau)^2}, \tag{26}$$

respectively. The lag-ℓ covariance of the process $\{Y_i\}$ is

$$\mathrm{Cov}(Y_i, Y_{i+\ell}) = \frac{1-\tau^2}{2\pi\lambda^2} \int_0^r \frac{z^{|\ell|+3/2}(r-z)^{1/2}}{(1-z)^3}\, dz \quad \text{for} \quad \ell = 0, \pm 1, \pm 2, \ldots, \tag{27}$$

where $r = 4\tau/(1+\tau)^2$ so that $0 < r < 1$; and the variance parameter is given by

$$\Omega_Y^2 = \frac{\tau^3\left(\tau^3 - 4\tau^2 + 5\tau + 2\right)}{\lambda^2\left(1-\tau\right)^4}. \tag{28}$$

The service rate of the in-control process is set to $\nu = 1$. To test different levels of dependence, we took the arrival rate $\lambda \in \{0.3, 0.6\}$ so that for the traffic intensity of the in-control system, we have $\tau \in \{0.3, 0.6\}$. We generate the monitored process $\{Y_i : i = 1, 2, \}$ based on the algorithm of Schmeiser and Song (1989) to so that the process is stationary with the steady-state properties (25)–(28). To generate shifted data, we first generate observations from an in-control process, and then add a constant to the observations. Therefore, the mean parameter changes, but variance parameters do not. Similar to AR(1) processes, the shift varies over 0, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3, and 4 in multiples of σ_Y. For the values $\tau = 0.3$ and $\tau = 0.6$ of the traffic intensity, we require the batch sizes $m = 2$ and $m = 10$, respectively, to achieve an approximate lag-one correlation of 0.5 for the batch means of the queue waiting time process. Similarly for the values $\tau = 0.3$ and $\tau = 0.6$ of the traffic intensity, we require the batch sizes $m = 11$ and $m = 55$, respectively, to achieve an approximate lag-one correlation of 0.1 for the batch means of the queue waiting time process.

As shown in Table 5, J&B and New CuSum achieve values of ARL_0 that are close to the target value 10,000. However, due to correlation and nonnormality of the monitored process $\{Y_i\}$, there is some deviation from the target ARL_0 for the distribution-free tabular CUSUM chart. Batching helps to reduce this deviation from the target value of ARL_0 with a small degradation in performance with respect to ARL_1. The performance of R&W is significantly degraded because of the large batch size required to achieve approximately i.i.d. *normal* batch means. Because of the nonnormality of Y_i as revealed in (25), the batch size m_1 that results in a lag-one correlation of approximately 0.1 is not large enough to achieve approximate normality of the batch means; and the actual value of ARL_0 deviates substantially from the target value of 10,000. For example, with $m_1 = 11$, the R&W procedure delivered $\text{ARL}_0 = 700$ when $\tau = 0.3$. To calibrate the R&W procedure in this situation, we increased the batch size m until the estimated ARL_0 was close to the target value. The resulting batch sizes are quite large—we must take $m = 300$ when $\tau = 0.3$; and we must take $m = 400$ when $\tau = 0.6$. Such large batch sizes cause catastrophic degradation in the performance of the R&W procedure in this test process.

TABLE 5. Two-Sided ARLs in Terms of Number of Raw Observations for an $M/M/1$ Queue.

τ	Shift	J&B	New CuSum	Dist.-Fr. CUSUM		R&W	
					$m = 2$	$m_1 = 11$	$m = 300$
0.3	0	10620	10374	8681	9236	700	9535
	0.25	1108	796	595	596		632
	0.5	554	393	231	238		300
	0.75	368	260	139	146		300
	1	276	196	99	105		300
	1.5	184	130	64	68		300
	2	138	97	47	50		300
	2.5	110	78	37	40		300
	3	92	65	31	33		300
	4	69	49	23	25		300
					$m = 10$	$m_1 = 55$	$m = 400$
0.6	0	11589	11259	14007	13504	2680	10274
	0.25	2380	1725	1893	1830		2381
	0.5	1185	847	735	746		655
	0.75	782	557	446	463		402
	1	583	414	318	337		400
	1.5	389	275	202	217		400
	2	290	205	148	161		400
	2.5	233	165	117	128		400
	3	194	136	97	107		400
	4	145	103	72	81		400

Conclusions and Recommendations

We presented a distribution-free CUSUM chart with the nonzero reference value for autocorrelated data. The experimental results strongly suggest that the proposed chart reacts more quickly to meaningful shifts than other existing distribution-free CUSUM charts. The new chart provides a simple way to determine the control limits, and it allows for the use of raw observations. To improve the accuracy of the setup for determining the control limits, batching can be used. The batch size required for this purpose is usually quite small, and a routinely applicable method for choosing a good batch size is provided. In terms of chart performance, our experiments demonstrate that if the monitored process is approximately Gaussian with small to moderate lag-one correlation, then the distribution-free tabular CUSUM chart generally outperforms the other existing distribution-free CUSUM-type charts and is competitive with the R&W chart. If the monitored process exhibits marked departures from normality or a pronounced dependency structure, then our experimental results indicate that the distribution-free tabular CUSUM chart significantly outperforms existing distribution-free SPC charts for autocorrelated data, including the R&W chart.

The chief limitation of the experimentation reported in this article is that it is based on the assumption that the marginal variance σ_Y^2 and the variance parameter Ω_Y^2 are known quantities. In many practical applications, the uncertainty about the values of these quantities is at least as great as the uncertainty about the value of the process mean μ; and thus extensive follow-up analysis and experimentation is required to evaluate the performance of the selected SPC procedures for monitoring shifts in the process mean μ when those procedures are augmented with appropriate variance-estimation procedures. This is the subject of ongoing research.

References

ALEXOPOULOS, C. AND GOLDSMAN, D. (2005). "Stationary Processes: Statistical Estimation". In Johnson, N. L., and Read, C. B., eds., *The Encyclopedia of Statistical Sciences*, 2nd ed. Wiley, New York.

BAGSHAW, M. AND JOHNSON, R. A. (1975). "The Effect of Serial Correlation on the

Performance of CUSUM Tests II". *Technometrics* 17, pp. 73–80.

Bickel, P. J. and Doksum, K. A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day, San Francisco, CA.

Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, New York.

Carlstein, E. (1986). "The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence". *The Annals of Statistics* 14, pp. 1171–1179.

Chien, C., Goldsman, D., and Melamed, B. (1997). "Large-Sample Results for Batch Means". *Management Science* 43, pp. 1288–1295.

Darling, D. A. and Siegert, A. J. F. (1953). "The First Passage Problem for a Continuous Markov Process". *Annals of Mathematical Statistics* 24, pp. 624–639.

Fuller, W. A. (1996). *Introduction to Statistical Time Series*, 2nd ed. John Wiley & Sons, New York.

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

Kim, S.-H., Alexopoulos, C., Goldsman, D., and Tsui, K.-L. (2005). "A New Model-Free CuSum Procedure for Autocorrelated Processes". Technical Report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Kim, S.-H., Nelson, B. L., and Wilson, J. R. (2005). "Some Almost-sure Convergence Properties Useful in Sequential Analysis". *Sequential Analysis* forthcoming. Available online via `ftp://ftp.ncsu.edu/pub/eos/pub/jwilson/kim05sqa.pdf` [accessed August 17, 2005].

Montgomery, D. C. (2001). *Introduction to Statistical Quality Control*, 4th ed. John Wiley & Sons, New York.

Park, Y. (2005). "A Statistical Process Control Approach for Network Intrusion Detection". Ph.D. Dissertation, School of Industrial and Systems Engineering, Atlanta, GA.

Runger, G. C. and Willemain, T. R. (1995). "Model-Based and Model-Free Control of Autocorrelated Processes". *Journal of Quality Technology* 27, pp. 283–292.

Schmeiser, B. W. and Song, W. T. (1989). "Inverse-transformation algorithms for some common stochastic processes". In MacNair, E. A., Musselman, K. J., and Heidelberger, P., eds., *Proceedings of the 1989 Winter Simulation conference*, pp. 490–496. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals.* Springer-Verlag, New York.

Steiger, N. M., Lada, E. K., Wilson, J. R., Joines, J. A., Alexopoulos, C., and Goldsman, D. (2005). "ASAP3: A Batch Means Procedure for Steady-State Simulation Output Analysis". *ACM Transactions on Modeling and Computer Simulation* 15, pp. 39–73.

Steiger, N. M., and Wilson, J. R. "Convergence Properties of the Batch Means Method for Simulation Output Analysis". *INFORMS Journal on Computing* 13, pp. 277–293.

Stuart, A. and Ord, J. K. (1994). *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory*, 6th ed. Edward Arnold, London.

Wagner, M. A. F., and Wilson, J. R. (1996). "Using Univariate Bézier Distributions to Model Simulation Input Processes". *IIE Transactions* 28, pp. 699–711.

Ye, N., Li, X., Chen, Q., Emran, S. M., and Xu, M. (2001). "Probabilistic Techniques for Intrusion Detection Based on Computer Audit Data". *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* 31, pp. 266–274.

Ye, N., Vilbert, S., and Chen, Q. (2003). "Computer Intrusion Detection through EWMA for Autocorrelated and Uncorrelated Data". *IEEE Transactions on Reliability* 52, pp. 75–82.