A CUSUM Chart for Monitoring a Proportion with Autocorrelated Binary Observations

SHABNAM MOUSA VI
Georgia State University, Atlanta, GA 30303, USA
Max Planck Institute for Human Development, Berlin 14195, Germany

MARION R. REYNOLDS, JR.
Virginia Tech, Blacksburg, VA 24061-0439, USA
A CUSUM Chart for Monitoring a Proportion with Autocorrelated Binary Observations

SHABNAM MOUSA VI
Georgia State University, Atlanta, GA 30303, USA
Max Planck Institute for Human Development, Berlin 14195, Germany

MARION R. REYNOLDS, JR.
Virginia Tech, Blacksburg, VA 24061-0439, USA

When traditional control charts are used to monitor a proportion $p$, it is assumed that the binary observations are independent. This paper investigates the problem of monitoring $p$ when there is a continuous stream of autocorrelated binary observations that follow a two-state Markov chain model with first-order dependence. It is shown that both the Shewhart $p$-chart and the most efficient chart for independent observations, the Bernoulli CUSUM chart, are not robust to autocorrelation, and that adjusting the control limits of these traditional charts to account for the autocorrelation is not an efficient approach. Here we construct a Markov binary CUSUM (MBCUSUM) chart based on a log-likelihood-ratio statistic and show that this chart can be well approximated by using a Markov chain model, for which exact properties are calculable. Numerical results show that the MBCUSUM chart will detect most increases in $p$ faster than competing charts. The effect of the size of the Phase I data set used in setting up the MBCUSUM chart is also investigated.

Key Words: Bernoulli CUSUM; Binary Data; Estimating Process Parameters; Markov Chain; Phase I Sample Size; Surveillance.

When control charts are used to monitor a process, the data obtained from the process will sometimes be binary observations that can take on only two values, 0 and 1. Binary observations may arise as the natural outcome of an inspection process or from categorizing some continuous measurements as conforming or nonconforming to some given standards. Although the application of control charts for binary observations now extends far beyond the traditional industrial setting (e.g., see Woodall (2006) for applications in health-care monitoring), for convenience, we use the traditional terms nondefective and defective for the values 0 and 1, respectively.

Most published work on control charts for binary observations assumes that the observations are independent. However, in recent decades, there has been increasing awareness that the observations from many processes are autocorrelated and that autocorrelation can have an adverse effect on the performance of control charts (Alwan and Roberts (1995)). See Broadbent (1958) and Deligonul and Mergen (1987) for investigations of the case of binary observations.

Most of the published work on control charts for autocorrelated binary observations assumes that the
observations can be modeled as a two-state Markov chain in which the probability of an observation being defective depends on the value of the previous observation (first-order dependence). See, e.g., Blatterman and Champ (1992), Champ et al. (1994), and Shepherd et al. (2007).

The main objective of this paper is to develop a CUSUM chart, called the Markov binary CUSUM (MBCUSUM) chart, for monitoring a process in which the observations are binary and follow the two-state Markov chain model. We consider the situation in which a continuous stream of binary observations is available for process monitoring (as with 100% inspection). It is assumed that these binary observations become available individually, so the control charts can be based on samples of $n = 1$. The MBCUSUM chart is based on a log-likelihood-ratio statistic derived from the two-state Markov chain model. This work can be considered to be an extension of the Reynolds and Stoumbos (1999) paper, which developed the Bernoulli CUSUM chart for a continuous stream of independent binary observations (see also Reynolds and Stoumbos (2000, 2001)).

Additional published work has considered monitoring problems related to what is investigated here. For example, Bhat and Lal (1990) and Tang and Cheong (2006) considered monitoring autocorrelated binary observations when samples, rather than a continuous stream of observations, are available from the process, and Niaki and Abbasi (2007) considered the situation in which there is autocorrelation and a defective item can have more than one defect.

We next define the two-state Markov chain model and the performance measures that will be used to evaluate control charts. We then show that control charts used for independent binary observations are not robust to autocorrelation, and thus establish the need for control charts that explicitly account for autocorrelation. Then we develop the MBCUSUM chart and show that, compared with other charts, it will provide faster detection of most shifts in $p$. Finally, we use an example to show how the MBCUSUM chart can be set up using a Phase I data set and investigate the effect of the size of the Phase I data set on the properties of the MBCUSUM chart.

The Two-State Markov Chain Model

Consider a sequence $X_1, X_2, X_3, \ldots$ of binary observations taking the values 0 and 1 (nondefective and defective, respectively). We refer to these observations as binary observations, rather than Bernoulli observations, because “Bernoulli” is usually associated with the case of independent observations. The first observation $X_1$ will be observed without knowing the value of a previous observation, so we assume that $X_1$ is a binary observation with $P(X_1 = 1) = p$ and $P(X_1 = 0) = 1 - p$. Once $X_1$ is observed, the remaining $X_2, X_3, X_4, \ldots$ can be generated using the two-state Markov chain model.

The transition probabilities, $p_{ij}, i, j = 0, 1$, for the two-state Markov chain model satisfy $p_{00} = 1 - p_{11}, i = 0, 1$, so this model can be characterized using only two parameters, $p_{00} = P(X_k = 1 | X_{k-1} = 0)$ and $p_{10} = P(X_k = 0 | X_{k-1} = 1)$ (see Bhat (1984) or Bhat and Lal (1990), where $p_{01}$ is labeled $a$ and $p_{10}$ is labeled $b$). The long-run proportion defective $p$ and the correlation coefficient $\rho$ between successive observations can be expressed as $p = P(X_k = 1) = p_{01}/(p_{01} + p_{10})$ and $\rho = 1 - (p_{01} + p_{10})$, respectively. For process monitoring applications, it seems more natural to us to parameterize the process in terms of $p$ and $\rho$ instead of $p_{01}$ and $p_{10}$. Then, for given values of $p$ and $\rho$, the transition probabilities are

\[
\begin{align*}
    p_{00} &= 1 - p(1 - \rho), \\
    p_{01} &= p(1 - \rho), \\
    p_{10} &= (1 - p)(1 - \rho), \\
    p_{11} &= 1 - (1 - p)(1 - \rho).
\end{align*}
\]

Let $p_0$ be the in-control value of $p$. We assume that the objective of process monitoring is to detect any change in the process that increases $p$ above $p_0$. When there is a shift in $p$, we assume that the values of both $p_{01}$ and $p_{10}$ shift in such a way that $\rho$ remains unchanged. Detecting a decrease in $p$ may also be of interest in some applications, but here we do not consider this problem.

Applying this model in practice usually requires that the in-control parameter values be estimated during a Phase I analysis when process data are collected for this purpose. Suppose that there are $N$ observations in the Phase I data set, so that the observed number of transitions between two states is $N - 1$. The maximum likelihood estimator of $p_{ij}$ is (Bhat and Lal (1990)) $\hat{p}_{ij} = N_{ij}/N_{01} + N_{11}$, where $N_{ij}$ is the number of transitions from state $i$ to state $j$ for $i, j = 0, 1$. Then the estimators of $p_0$ and $\rho$ are

\[
\begin{align*}
    \hat{p}_0 &= \frac{\hat{p}_{01}}{\hat{p}_{01} + \hat{p}_{10}} \quad \text{and} \quad \hat{\rho} = 1 - (\hat{p}_{01} + \hat{p}_{10}).
\end{align*}
\]
A practical concern is the effect of errors in estimating these parameters, so we address this issue later in the paper.

**Performance Measures for Control Charts**

Control charts are usually evaluated using the average run length (ARL), defined as the expected number of samples to signal. The ARL is not an appropriate performance measure when comparing control charts based on different sample sizes, so instead we use the average number of observations to signal (ANOS), defined as the expected number of observations from the start of process monitoring until a signal by the control chart (Reynolds and Stoumbos (2000, 2001)).

When the process is in-control \((p = p_0)\), we want the ANOS to be large so that the frequency of false alarms is low. If \(p\) increases to some value above \(p_0\), then we need a measure of how quickly this increase is detected. Some control charts, such as CUSUM charts, accumulate information over time, so we must account for the fact that the control statistics of these charts may not be at their starting values when the increase in \(p\) occurs. Thus, as a measure of the detection time, we use the steady state ANOS (SSANOS), which is based on the assumption that the distribution of the control statistic at the time that the increase in \(p\) occurs is the steady-state distribution of this statistic, conditional on no false alarms.

**Traditional Control Charts for Monitoring \(p\)**

The traditional control chart for monitoring \(p\) is the Shewhart \(p\)-chart (see Woodall (1997) for a general review of control charts for monitoring \(p\)), which is based on the assumption that the observations are independent. To apply this chart in the case of a continuous stream of binary observations, the stream of observations would be partitioned into segments of \(n\) observations (these segments would then constitute the samples from the process). If \(S_i\) is the number of defectives in the \(i\)th segment, then the Shewhart \(p\)-chart would signal that \(p\) has increased if \(S_i/n\) is above an upper control limit, which is equivalent to signaling if \(S_i \geq h_S\), for some constant \(h_S\). Instead of the traditional “three-sigma” control limits, here we choose \(h_S\) to give a desired value of the in-control ANOS when the observations are independent, \(S_i\) has a binomial distribution, but this does not hold when there is autocorrelation.

When the Shewhart \(p\)-chart is being used, the shift in \(p\) may occur in the middle of a sample, so the SSANOS of this chart is based on the assumption that the position of the shift within a sample of \(n\) observations has a uniform distribution. The ANOS and SSANOS were computed by using a Markov chain model (see the Appendix).

The Bernoulli CUSUM chart was developed by Reynolds and Stoumbos (1999) for monitoring \(p\) when there is a continuous stream of independent binary observations, each treated as a sample of \(n = 1\). The log-likelihood-ratio-based Bernoulli CUSUM control statistic at observation \(k\) is given by

\[
B_k = \max\{0, B_{k-1}\} + (X_k - \gamma_B), \quad k = 1, 2, \ldots, (2)
\]

where \(\gamma_B = -\ln((1 - p_1)/(1 - p_0))/\ln((p_1(1 - p_0))/(p_0(1 - p_1)))\) and \(p_1 > p_0\) is a value of \(p\) that should be detected quickly. This chart signals if \(B_k \geq h_B\) for some upper control limit \(h_B\).

The chart parameter \(p_1\) (which, for a given value of \(p_0\), determines the value of \(\gamma_B\)) can be used as a tuning parameter for the Bernoulli CUSUM chart. Choosing \(p_1\) to be close to \(p_0\) will make the chart particularly sensitive to small increases in \(p\), while choosing a larger value for \(p_1\) will make the chart sensitive to larger increases in \(p\). In fact, a particular choice of \(p_1\) will make the CUSUM chart optimal for detecting an increase in \(p\) from \(p_0\) to \(p = p_1\), in the sense that the ANOS at \(p = p_1\) is minimized subject to a specified value for the in-control ANOS. However, the SSANOS at \(p = p_1\) will not be minimized; in terms of the SSANOS, the chart will be optimal for a slightly different value of \(p\). We believe that the SSANOS is the most reasonable single measure of out-of-control performance, so the precise specification of the tuning parameter \(p_1\) is not critical in applications.

Reynolds and Stoumbos (1999) showed that, if \(p_0\) is not large, then for a given \(p_1\), a very slight adjustment of \(p_1\) can be made so that \(\gamma_B = 1/m\), where \(m\) is a positive integer. If \(\gamma_B = 1/m\), then \(B_k\) will be a lattice random variable for which the possible values are integer multiples of \(1/m\), and this will allow the Bernoulli CUSUM chart to be modeled exactly as a Markov chain, which in turn allows for the exact computation of the ANOS and SSANOS.

The problem of monitoring \(p\) formulated as observing independent Bernoulli observations can be equivalently formulated as one of observing the number of nondefectives between defectives and using the
geometric distribution. For example, Bourke (1991) investigated a geometric CUSUM chart based on the geometric observations. The values of the sequence of Bernoulli observations determine the sequence of geometric observations, and vice versa, so the two sequences contain the same information about the process. Reynolds and Stoumbos (1999) showed that the geometric CUSUM chart is equivalent to a Bernoulli CUSUM chart that starts with a headstart, so there is no need to consider the geometric CUSUM chart here separately from the Bernoulli CUSUM chart.

For the case of independent observations, Sego et al. (2007) and Joner et al. (2008) evaluated the performance of the Bernoulli CUSUM chart relative to the performance of some surveillance schemes traditionally used in health-care settings. They found that the Bernoulli CUSUM chart has better performance than these schemes in almost all cases.

### Robustness of Traditional Control Charts to Autocorrelation

We now investigate the robustness of standard control charts (designed under the assumption of independent binary observations) when the data actually follow the two-state Markov chain model with first-order dependence.

Consider the situation in which a Shewhart \( p \)-chart based on samples of \( n = 100 \) is used to monitor a process with \( p_0 = .010 \). If \( h_S = 5 \) for this control chart, then the in-control ANOS will be 29134.8 when there is no autocorrelation (this corresponds to 291.3 samples when \( n = 100 \)). The column labeled [1] in Table 1 gives the in-control ANOS of this chart for some values of \( \rho > 0 \). When \( p \) remains at \( p_0 \) and \( \rho \) increases, the values of \( p_{01} \) and \( p_{11} \) also change, so the values of \( p_{01} \) and \( p_{11} \) are also given in Table 1 for easy reference. The number of states used in modeling the Shewhart chart as a Markov chain is given at the bottom of Table 1.

Columns [2] and [3] of Table 1 give in-control ANOS values for the Bernoulli CUSUM chart with \( p_1 = .025 \) and \( p_1 = .040 \), respectively. The values of \( h_B \) were adjusted so that the in-control ANOS would be very close to the value 29134.8 for the Shewhart chart when \( \rho = 0 \).

From Table 1, we see that neither the Shewhart chart nor the Bernoulli CUSUM chart is robust to autocorrelation. As \( \rho \) increases above 0, the in-control ANOS values increase.

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( p_{01} )</th>
<th>( n = )</th>
<th>Shewhart</th>
<th>Bernoulli CUSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_{11} )</td>
<td>( h = 5 )</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( .00 )</td>
<td>.01</td>
<td>.01</td>
<td>29134.8</td>
<td>29248.6</td>
</tr>
<tr>
<td>( .05 )</td>
<td>.0095</td>
<td>.0595</td>
<td>16956.9</td>
<td>18464.7</td>
</tr>
<tr>
<td>( .10 )</td>
<td>.0090</td>
<td>.1090</td>
<td>11200.4</td>
<td>12661.0</td>
</tr>
<tr>
<td>( .15 )</td>
<td>.0085</td>
<td>.1585</td>
<td>7987.2</td>
<td>9204.0</td>
</tr>
<tr>
<td>( .20 )</td>
<td>.0080</td>
<td>.2080</td>
<td>6000.4</td>
<td>6988.4</td>
</tr>
<tr>
<td>( .25 )</td>
<td>.0075</td>
<td>.2575</td>
<td>4682.8</td>
<td>5487.9</td>
</tr>
<tr>
<td>( .30 )</td>
<td>.0070</td>
<td>.3070</td>
<td>3763.3</td>
<td>4427.0</td>
</tr>
<tr>
<td>( .35 )</td>
<td>.0065</td>
<td>.3565</td>
<td>3096.7</td>
<td>3651.1</td>
</tr>
<tr>
<td>( .40 )</td>
<td>.0060</td>
<td>.4060</td>
<td>2599.0</td>
<td>3068.0</td>
</tr>
<tr>
<td>( .45 )</td>
<td>.0055</td>
<td>.4555</td>
<td>2219.3</td>
<td>2620.4</td>
</tr>
<tr>
<td>( .50 )</td>
<td>.0050</td>
<td>.5050</td>
<td>1925.4</td>
<td>2271.3</td>
</tr>
<tr>
<td>Number of states</td>
<td></td>
<td></td>
<td>1200</td>
<td>640</td>
</tr>
</tbody>
</table>

Columns [2] and [3] of Table 1 give in-control ANOS values for the Bernoulli CUSUM chart with \( p_1 = .025 \) and \( p_1 = .040 \), respectively. The values of \( h_B \) were adjusted so that the in-control ANOS would be very close to the value 29134.8 for the Shewhart chart when \( \rho = 0 \).

From Table 1, we see that neither the Shewhart chart nor the Bernoulli CUSUM chart is robust to autocorrelation. As \( \rho \) increases above 0, the in-control ANOS values increase.
ANOS drops rapidly from what would be expected for independent observations, and this, of course, means that false alarms will occur much more frequently than expected. Another example with means that false alarms will occur much more frequently for independent observations, and this, of course, shows that this chart has better ability to detect shifts than the traditional charts, even with adjusted control limits.

A CUSUM Chart for Autocorrelated Data

The proposed MBCUSUM chart is based on the log-likelihood-ratio statistics for an increase in $p_0$ to $p_1$ in the two-state Markov chain model. The joint density of $X_1, X_2, \ldots, X_k$ can be written as

$$f(x_1, x_2, \ldots, x_k \mid p) = f(x_1 \mid p) \prod_{i=2}^{k} f(x_i \mid x_{i-1}, p),$$

so it follows that the terms that we need to use in the CUSUM control statistic are

$$L_k = \begin{cases} \ln f(x_1 \mid p_0) - \ln f(x_1 \mid p_1), & k = 1, \\ \ln f(x_k \mid x_{k-1}, p_1) - \ln f(x_1 \mid x_{k-1}, p_0), & k = 2, 3, \ldots \end{cases}$$

Using the Bernoulli distribution of $X_1$ gives

$$L_1 = x_1 \ln \frac{p_1}{p_0} + (1 - x_1) \ln \frac{1 - p_1}{1 - p_0}. \quad (3)$$

Now for $k \geq 2$,

$$f(x_k \mid x_{k-1}) = p_0^{(1-x_{k-1})(1-x_k)} p_1^{(1-x_{k-1})x_k} x_{k-1}^{1-x_k} x_{k-1}^{1-x_k},$$

and this gives

$$L_k = (1 - x_{k-1})(1 - x_k) l_{00} + (1 - x_{k-1})x_k l_{01} + x_{k-1}(1 - x_k) l_{10} + x_{k-1}x_k l_{11},$$

where

$$\begin{align*}
l_{00} &= \ln \frac{1 - p_1(1 - \rho)}{1 - p_0(1 - \rho)}, \\
l_{01} &= \ln \frac{p_1}{p_0}, \\
l_{10} &= \ln \frac{1 - p_1}{1 - p_0};
\end{align*}$$

$$l_{11} = \ln \frac{1 - (1 - p_1)(1 - \rho)}{1 - (1 - p_0)(1 - \rho)}.$$ 

Thus, we see that, for $k \geq 2$,

$$L_k = \begin{cases} l_{00} & \text{if } x_{k-1} = 0 \text{ and } x_k = 0, \\ l_{01} & \text{if } x_{k-1} = 0 \text{ and } x_k = 1, \\ l_{10} & \text{if } x_{k-1} = 1 \text{ and } x_k = 0, \\ l_{11} & \text{if } x_{k-1} = 1 \text{ and } x_k = 1. \end{cases} \quad (4)$$

The control statistic for the MBCUSUM chart for $k = 1, 2, \ldots$ is

$$C_k = \max\{0, C_{k-1}\} + L_k,$$

where $C_0 = 0$. A signal is given if $C_k \geq h_C$. Note that, when $\rho = 0$, we get $l_{00} = l_{10}$ and $l_{01} = l_{11}$ and the MBCUSUM reduces to the Bernoulli CUSUM (Equation (2)) when we divide $C_k$ by $l_{01} - l_{10}$.

Recall that properties of the Bernoulli CUSUM chart can be evaluated by using a slight modification of the problem so that the CUSUM control statistic is a lattice random variable with possible values that are integer multiples of $1/m$, where $m$ is a positive integer. This allows the Bernoulli CUSUM to be modeled as a Markov chain. For the MBCUSUM chart, we adopt a similar strategy and get an approximate MBCUSUM chart by approximating $L_k$ by a random variable for which possible values are integer multiples of a constant. The approach that we used to approximate $L_k$ is to obtain the integer $m = \text{nint}\{1/l_{00}\}$, where $\text{nint}()$ indicates the nearest integer value. Then $L_k$ is approximated by a new statistic, say $L_k^*$, for which the possible values are integer multiples of $1/m$. In particular, for $k \geq 2$,

$$L_k^* = \begin{cases} \text{nint}(l_{00} m)/m & \text{if } x_{k-1} = 0 \text{ and } x_k = 0, \\ \text{nint}(l_{01} m)/m & \text{if } x_{k-1} = 0 \text{ and } x_k = 1, \\ \text{nint}(l_{10} m)/m & \text{if } x_{k-1} = 1 \text{ and } x_k = 0, \\ \text{nint}(l_{11} m)/m & \text{if } x_{k-1} = 1 \text{ and } x_k = 1. \end{cases}$$

Now for the first observation, $L_1 = l_{10}$ when $x_1 = 0$ and $L_1 = l_{01}$ when $x_1 = 1$, so $L_1^* = l_{10}$ when $x_1 = 0$ and $L_1^* = l_{01}$ when $x_1 = 1$. The approximate MBCUSUM control statistic then is

$$C_k^* = \max\{0, C_{k-1}^*\} + L_k^*, \quad k = 1, 2, \ldots,$$

where $C_0^* = 0$. A signal is given if $C_k^* \geq h_C$, where $h_C$ now is taken as an integer multiple of $1/m$. For more details, see Mousavi and Reynolds (2008).

Comparisons of the MBCUSUM with Other Control Charts

In this section, we compare the performance of the exact and approximate MBCUSUM chart against the traditional Shewhart $p$ chart and the Bernoulli
TABLE 2. ANOS and SSANOS Values for Shewhart and CUSUM Charts for Detecting Increases in $p$
when $p_0 = .010$ and $p_1 = .025$ or .040, and $p = .05$

<table>
<thead>
<tr>
<th>$p$</th>
<th>$p_0$</th>
<th>$p_1$</th>
<th>$n =$</th>
<th>Shewhart Standard</th>
<th>Curtailed Bernoulli</th>
<th>CUSUM, $p_1 = .025$ Bernoulli Exact</th>
<th>CUSUM, $p_1 = .040$ Bernoulli Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>.010</td>
<td>.0095</td>
<td>.0595</td>
<td>100</td>
<td>16956.9</td>
<td>16935.5</td>
<td>16977.5</td>
<td>16869.6</td>
</tr>
<tr>
<td>.015</td>
<td>.0142</td>
<td>0.643</td>
<td>4226.2</td>
<td>4210.2</td>
<td></td>
<td>2351.4</td>
<td>2195.9</td>
</tr>
<tr>
<td>.020</td>
<td>.0190</td>
<td>0.690</td>
<td>1678.5</td>
<td>1658.3</td>
<td></td>
<td>848.3</td>
<td>796.8</td>
</tr>
<tr>
<td>.025</td>
<td>.0238</td>
<td>0.738</td>
<td>870.6</td>
<td>848.3</td>
<td></td>
<td>473.3</td>
<td>448.0</td>
</tr>
<tr>
<td>.030</td>
<td>.0285</td>
<td>0.785</td>
<td>536.6</td>
<td>512.8</td>
<td></td>
<td>322.3</td>
<td>306.7</td>
</tr>
<tr>
<td>.040</td>
<td>.0380</td>
<td>0.880</td>
<td>281.4</td>
<td>254.9</td>
<td></td>
<td>195.2</td>
<td>187.3</td>
</tr>
<tr>
<td>.050</td>
<td>.0475</td>
<td>0.975</td>
<td>190.3</td>
<td>161.2</td>
<td></td>
<td>139.8</td>
<td>135.2</td>
</tr>
<tr>
<td>.070</td>
<td>.0665</td>
<td>1.165</td>
<td>125.4</td>
<td>91.7</td>
<td></td>
<td>89.1</td>
<td>87.5</td>
</tr>
<tr>
<td>.100</td>
<td>.0950</td>
<td>1.450</td>
<td>95.6</td>
<td>56.6</td>
<td></td>
<td>57.7</td>
<td>57.9</td>
</tr>
<tr>
<td>.200</td>
<td>.1900</td>
<td>2.400</td>
<td>70.5</td>
<td>25.4</td>
<td></td>
<td>26.7</td>
<td>28.4</td>
</tr>
<tr>
<td>.300</td>
<td>.2850</td>
<td>3.350</td>
<td>62.7</td>
<td>16.3</td>
<td></td>
<td>17.5</td>
<td>19.9</td>
</tr>
<tr>
<td>.400</td>
<td>.3800</td>
<td>4.300</td>
<td>58.8</td>
<td>12.0</td>
<td></td>
<td>13.1</td>
<td>16.0</td>
</tr>
<tr>
<td>.500</td>
<td>.4750</td>
<td>5.250</td>
<td>56.5</td>
<td>9.5</td>
<td></td>
<td>10.4</td>
<td>14.0</td>
</tr>
<tr>
<td>.700</td>
<td>.6650</td>
<td>7.150</td>
<td>53.9</td>
<td>6.7</td>
<td></td>
<td>7.4</td>
<td>12.4</td>
</tr>
<tr>
<td>.900</td>
<td>.8550</td>
<td>9.050</td>
<td>52.5</td>
<td>5.2</td>
<td></td>
<td>5.7</td>
<td>13.0</td>
</tr>
</tbody>
</table>

$h_S = 5$  
Number of states 1200

CUSUM chart and against a chart developed by Shepherd et al. (2007) for autocorrelated data from the two-state Markov chain model. The Shewhart $p$-chart and Bernoulli CUSUM chart are not robust to autocorrelation (when designed assuming independent observations), so the objective here is to see how well these charts compare with the MBCUSUM chart when the control limits of all charts are adjusted to give the same in-control ANOS in the presence of autocorrelation.

The Traditional Control Charts

When comparing different control charts, the discreteness of the distributions usually prevents us from obtaining exactly the same in-control ANOS for all of the charts. The Shewhart chart has only a very few possible values for the in-control ANOS, so we first chose the control limit for the Shewhart chart, and then found the control limits of the CUSUM charts to obtain a close match to the in-control ANOS of the Shewhart chart.

Consider first a process with $p_0 = .010$ and $p = .05$. A Shewhart chart based on samples of $n = 100$ will have an in-control ANOS of 16956.9 if the control limit is $h_S = 5$. Assume, for purposes of this example, that this is an acceptable in-control ANOS. The column labeled [1] in Table 2 gives out-of-control SSANOS values for this chart for various values of $p > p_0$ (all assuming that $p = .05$).

The Shewhart $p$-chart is based on dividing the continuous stream of observations into samples of $n = 100$, so it is very likely that any shift in $p$ will occur somewhere in the middle of a sample. A signal can be given only at the end of a segment, so the expected time from the shift to the signal (the SSANOS) can be less than 100, but it will always be at least 50, even for very high values of $p$.

It seems clear that the performance of the Shewhart chart could be improved by using curtailed sampling in which a signal would be given as soon as five defectives have been found in a sample (without waiting until the end of the sample). Column [2] in Table 2 gives the in-control ANOS and out-of-control SSANOS values for the Shewhart chart with curtailed sampling. We see that using curtailed sam-
Sampling produces a slight reduction in the in-control ANOS from 16956.9 to 16935.5, but gives a dramatic reduction in the SSANOS for very large shifts in \( p \). Although it is clear that curtailed sampling can be quite beneficial if there is a large shift in \( p \), curtailed sampling seems to have been rarely used in applications.

In Table 2, columns [3], [4], and [5], respectively, give the in-control ANOS and out-of-control SSANOS values for the Bernoulli CUSUM chart, the exact MBCUSUM chart, and the approximate MB-CUSUM chart when \( p_1 = .025 \). Columns [6], [7], and [8] correspond to the case of \( p_1 = .040 \). Here \( h_B \) and \( h_C \) for the CUSUM charts have been adjusted to give an in-control ANOS approximately the same as for the Shewhart chart. The ANOS and SSANOS values for the exact MBCUSUM were obtained by simulation using 100 million simulation runs. When the ANOS and SSANOS values of a control chart were obtained by using a Markov chain model, the number of states used is given at the bottom of the table.

In Table 2, we see that the exact MBCUSUM is very slightly better than the approximate MB-CUSUM for small shifts in \( p \), but the reverse is true for large shifts. However, the difference between the exact and approximate MBCUSUM charts is so small that the approximate MBCUSUM can be used instead of the exact MBCUSUM with negligible effect on the SSANOS performance. We also see that the MBCUSUM chart has better performance than the Bernoulli CUSUM chart except for some very large shifts in \( p \) (these very large shifts were included in the table to show how the charts perform in extreme situations). In most applications, the primary interest would likely be in the values of \( p \) for which the MB-CUSUM has better performance than the Bernoulli CUSUM, so we conclude that the MBCUSUM is a better choice when there is autocorrelation.

Comparing the CUSUM charts to the Shewhart charts in Table 2 shows that the CUSUM charts have much better performance than the Shewhart chart for small shifts in \( p \). The CUSUM charts also have much better performance than the Shewhart chart for large shifts in \( p \) unless curtailed sampling is used in the Shewhart chart. Even when curtailed sampling is used in the Shewhart chart, the Shewhart chart is uniformly worse than the Bernoulli CUSUM chart designed with \( p = .040 \).

Now consider the situation with the same value of \( p_0 = .010 \) but a much higher correlation coefficient, \( \rho = .20 \). If a Shewhart chart with \( n = 100 \) and a control limit of \( h_S = 5 \) is used, then the in-control ANOS of this chart is only 6000.4 (see Table 1). To obtain a larger in-control ANOS, consider a control limit of \( h_S = 6 \), which gives an in-control ANOS of 16890.0. Table 3 gives in-control ANOS and out-of-control SSANOS values for Shewhart and CUSUM charts for this situation, where \( h_B \) and \( h_C \) have been adjusted to give an in-control ANOS approximately the same as for the Shewhart chart. The structure of Table 3 is the same as for Table 2, and the basic conclusions are similar.

The in-control ANOS values in Table 3 are close to those of Table 2, so comparisons can be made between the case of \( \rho = .05 \) in Table 2 and the case of higher correlation (\( \rho = .20 \)) in Table 3. Comparing Tables 2 and 3 shows that detecting a given shift in \( p \) is harder (the SSANOS is larger) when the correlation is higher.

We conclude that, even if the autocorrelation is taken into account in designing the Bernoulli CUSUM chart or the Shewhart \( p \)-chart, the MB-CUSUM chart gives better overall performance in detecting increases in \( p \).

Another Scheme for Autocorrelated Data

For the case in which the binary observations follow the two-state Markov chain model, Shepherd et al. (2007) proposed two control charts based on the number of nondefectives between defectives. Let \( Y_1 \) be the number of nondefectives before the first defective, and let \( Y_j \) be the number of nondefectives between defectives \( j-1 \) and \( j \), for \( i = 2, 3, \ldots \). Shepherd et al. (2007) derived various properties of \( Y_1, Y_2, \ldots \), and showed that these random variables are independent. The distribution of \( Y_1 \) is different from \( Y_j \), \( j \geq 2 \) because the value of the observation before \( X_1 \) is unknown.

The second control chart proposed by Shepherd et al. (2007) seems to have the better properties, so we consider this chart here and refer to it as the SCRF chart (for Shepherd, Champ, Rigdon, and Fuller). The SCRF chart signals if two consecutive values of \( Y_j \) fall below a lower control limit, \( h_{SCRF} \). Shepherd et al. (2007) actually suggested using two control limits, one for \( Y_1 \) and the other for \( Y_j, j \geq 2 \), because the distributions are a bit different. However, the distribution of \( Y_1 \) has little effect on the SSANOS because the shift in \( p \) is assumed to occur after the process has reached its conditional steady-state distribution.
TABLE 3. ANOS and SSANOS Values for Shewhart and CUSUM Charts for Detecting Increases in p when \( p_0 = .010 \) and \( p_1 = .025 \) or .040, and \( \rho = .20 \)

<table>
<thead>
<tr>
<th>( p )</th>
<th>( p_{01} )</th>
<th>( p_{11} )</th>
<th>( n = 100 )</th>
<th>( n = 100 )</th>
<th>( n = 100 )</th>
<th>( n = 100 )</th>
<th>( n = 100 )</th>
<th>( n = 100 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( .010 )</td>
<td>.0080</td>
<td>.2080</td>
<td>16890.0</td>
<td>16863.3</td>
<td>16830.1</td>
<td>16821.1</td>
<td>16814.2</td>
<td>16815.9</td>
</tr>
<tr>
<td>( .015 )</td>
<td>.0120</td>
<td>.2120</td>
<td>5509.6</td>
<td>5493.1</td>
<td>2956.0</td>
<td>2394.3</td>
<td>2391.6</td>
<td>3989.9</td>
</tr>
<tr>
<td>( .020 )</td>
<td>.0160</td>
<td>.2160</td>
<td>2476.1</td>
<td>2454.8</td>
<td>1114.6</td>
<td>896.9</td>
<td>896.3</td>
<td>1499.4</td>
</tr>
<tr>
<td>( .025 )</td>
<td>.0200</td>
<td>.2200</td>
<td>1351.7</td>
<td>1328.4</td>
<td>625.5</td>
<td>509.5</td>
<td>509.3</td>
<td>773.1</td>
</tr>
<tr>
<td>( .030 )</td>
<td>.0240</td>
<td>.2240</td>
<td>842.7</td>
<td>818.2</td>
<td>426.0</td>
<td>350.3</td>
<td>350.3</td>
<td>487.6</td>
</tr>
<tr>
<td>( .040 )</td>
<td>.0320</td>
<td>.2320</td>
<td>427.2</td>
<td>400.8</td>
<td>257.6</td>
<td>214.8</td>
<td>214.8</td>
<td>268.3</td>
</tr>
<tr>
<td>( .050 )</td>
<td>.0400</td>
<td>.2400</td>
<td>272.3</td>
<td>244.2</td>
<td>184.2</td>
<td>155.4</td>
<td>155.4</td>
<td>182.8</td>
</tr>
<tr>
<td>( .070 )</td>
<td>.0560</td>
<td>.2560</td>
<td>160.9</td>
<td>129.3</td>
<td>117.3</td>
<td>100.8</td>
<td>100.8</td>
<td>111.1</td>
</tr>
<tr>
<td>( .100 )</td>
<td>.0800</td>
<td>.2800</td>
<td>77.1</td>
<td>71.9</td>
<td>77.1</td>
<td>66.9</td>
<td>66.9</td>
<td>70.0</td>
</tr>
<tr>
<td>( .200 )</td>
<td>.1600</td>
<td>.3600</td>
<td>67.2</td>
<td>20.9</td>
<td>22.7</td>
<td>24.2</td>
<td>24.1</td>
<td>19.9</td>
</tr>
<tr>
<td>( .400 )</td>
<td>.3200</td>
<td>.5200</td>
<td>62.3</td>
<td>15.3</td>
<td>16.6</td>
<td>20.2</td>
<td>20.2</td>
<td>14.6</td>
</tr>
<tr>
<td>( .500 )</td>
<td>.4000</td>
<td>.6000</td>
<td>59.4</td>
<td>12.1</td>
<td>13.1</td>
<td>18.4</td>
<td>18.3</td>
<td>11.6</td>
</tr>
<tr>
<td>( .700 )</td>
<td>.5600</td>
<td>.7600</td>
<td>56.0</td>
<td>8.5</td>
<td>9.3</td>
<td>18.5</td>
<td>18.3</td>
<td>8.2</td>
</tr>
<tr>
<td>( .900 )</td>
<td>.7200</td>
<td>.9200</td>
<td>54.2</td>
<td>6.6</td>
<td>7.2</td>
<td>27.5</td>
<td>26.6</td>
<td>6.4</td>
</tr>
</tbody>
</table>

\[ h = 6 \]

Number of states 1400 1400 814 682 506 414

so we simplify the specification of this chart by using one control limit.

The SCRF chart is actually a special case of what is usually called the “sets method” originally proposed by Chen (1978) in the context of health-care monitoring. With the sets method, a signal is given if a specified number of consecutive values of \( Y_j \) fall below \( h \). Previous work on the sets method, however, has been for the case of independent observations. Sego et al. (2007) recently did a comprehensive evaluation of the sets method, some variations of the sets method, and the Bernoulli CUSUM chart for the case of independent observations. The conclusion was that the Bernoulli CUSUM chart almost uniformly outperforms the sets method and its variations.

The SCRF chart seems to be most effective when \( p_0 \) is small, so to compare the SCRF chart with the MBCUSUM chart, we considered the case of \( p_0 = .001 \). We chose four SCRF charts and then adjusted \( h_C \) for the approximate MBCUSUM chart to closely match the in-control ANOS of the SCRF charts. In particular, for \( \rho = .05 \), we used \( h_{SCRF} = 50 \) and 200, and for \( \rho = .20 \), we used \( h_{SCRF} = 10 \) and 100. The SCRF chart seems to be particularly effective for detecting large shifts in \( p \), so we used \( p_1 = .008 \) in the MBCUSUM charts. The ANOS and SSANOS values for the four cases are given in Table 4.

The SSANOS values in Table 4 show that the MBCUSUM chart has better performance than the SCRF chart, except for very large shifts. Shifts that are very large, relative to \( p_0 = .001 \), would likely be of little concern in most applications, but were included here just to show how the charts perform in extreme situations. The SCRF chart is not dramatically better than the MBCUSUM for these large shifts, but for smaller values of \( p \), which are presumably of more interest, the MBCUSUM can be dramatically better than the SCRF chart.

Example

To illustrate the process of estimating process parameters and applying the MBCUSUM chart, consider a Phase I period in which \( N = 10,000 \) binary observations have been obtained. These Phase I
TABLE 4. ANOS and SSANOS Values for the SCRF and MBCUSUM Charts for Detecting Increases in $p$

<table>
<thead>
<tr>
<th>$p$</th>
<th>ANOS SCRF</th>
<th>Approx SCRF</th>
<th>ANOS MBCUSUM</th>
<th>Approx MBCUSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>0.0009</td>
<td>0.0510</td>
<td>26641.8</td>
<td>26656.6</td>
</tr>
<tr>
<td>0.002</td>
<td>0.0019</td>
<td>0.0519</td>
<td>5826.6</td>
<td>4975.5</td>
</tr>
<tr>
<td>0.003</td>
<td>0.0029</td>
<td>0.0529</td>
<td>2515.4</td>
<td>2107.0</td>
</tr>
<tr>
<td>0.004</td>
<td>0.0038</td>
<td>0.0538</td>
<td>965.9</td>
<td>505.4</td>
</tr>
<tr>
<td>0.005</td>
<td>0.0047</td>
<td>0.0547</td>
<td>454.7</td>
<td>319.4</td>
</tr>
<tr>
<td>0.006</td>
<td>0.0057</td>
<td>0.0557</td>
<td>34.2</td>
<td>31.9</td>
</tr>
<tr>
<td>0.007</td>
<td>0.0067</td>
<td>0.0566</td>
<td>3.2</td>
<td>3.1</td>
</tr>
<tr>
<td>0.008</td>
<td>0.0076</td>
<td>0.0576</td>
<td>2.3</td>
<td>2.2</td>
</tr>
<tr>
<td>0.009</td>
<td>0.0086</td>
<td>0.0586</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0096</td>
<td>0.0596</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.02</td>
<td>0.0196</td>
<td>0.0696</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.03</td>
<td>0.0296</td>
<td>0.0796</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.04</td>
<td>0.0396</td>
<td>0.0896</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0496</td>
<td>0.0996</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.06</td>
<td>0.0596</td>
<td>0.1096</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.07</td>
<td>0.0696</td>
<td>0.1196</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.08</td>
<td>0.0796</td>
<td>0.1296</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.09</td>
<td>0.0896</td>
<td>0.1396</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0996</td>
<td>0.1496</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1996</td>
<td>0.2496</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2996</td>
<td>0.3496</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3996</td>
<td>0.4496</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4996</td>
<td>0.5496</td>
<td>1.3</td>
<td>1.2</td>
</tr>
</tbody>
</table>

In practice, the values of $p_0$ and $\rho$ would usually be unknown, so we illustrate the process of estimating these parameters using Equation (1). Figure 1 shows the transitions for the 10,000 observations according to whether $X_k = (1,0)$, $(0,0)$, $(1,1)$, or $(0,1)$, respectively. These representations for $X_k$ are in the order corresponding to increasing values of the likelihood ratio $L_k$. The vast majority of the transitions in the Phase I data set are from 0 to 0 so these values plot as what appears to be a solid horizontal line at (0,0) in Figure 1. The transitions shown in Figure 1 result in $N_{00} = 9806$, $N_{01} = 84$, $N_{10} = 84$, and $N_{11} = 25$, so this gives $\hat{p}_0 = 0.010901$ and $\hat{\rho} = 0.220864$. These estimates differ a bit from the true values of $p_0$ and $\rho$ (the values used to simulate the data).
control ANOS of approximately 16900. If $\hat{p}_0 = 0.010901$ and $\hat{\rho} = 0.220864$ are used for $p_0$ and $\rho$ in the Markov chain model for the approximate MB-CUSUM chart, then we find that $h_C = 194/38 = 5.1053$ will give a nominal in-control ANOS of 16819.9.

Figure 2 is a plot of the approximate MBCUSUM chart for the Phase I data, with the control limit at $h_C = 5.1053$. The plot in Figure 2 shows no point outside of the control limit, so this does not provide any evidence of an unstable process in Phase I.

Once the unknown process parameters have been estimated and the control chart has been designed, the control chart can be used for monitoring the process in real time as more observations are obtained (this is Phase II). To illustrate process monitoring in Phase II, 1000 additional observations were simulated. The first 500 were in-control observations with $p = p_0 = 0.01$ and $\rho = 0.2$, while the last 500 had $p = p_1 = 0.04$ and $\rho = 0.2$, corresponding to an out-of-control situation. These 1000 observations are shown in Figure 3 using the same plotting convention as in Figure 1. We see that there seems to be an increase in the $(0, 1)$ transitions in the last 500 observations (and a corresponding increase in the $(1, 0)$ transitions).

The approximate MBCUSUM chart for the Phase II observations is plotted in Figure 4. This CUSUM plot clearly shows the increase in $p$ that occurred after observation 500, and the plot indicates reasonably well where the increase in $p$ occurred. In practice, of course, some action should be taken when the CUSUM statistic first crosses the control limit.

The Effect of Estimated Process Parameters

The approximate MBCUSUM chart in the example was designed using estimated parameter values of $\hat{p} = 0.010901$ and $\hat{\rho} = 0.220864$, as would usually have to be done in practice, and the in-control ANOS was determined by acting as if these estimates are the true process parameters. An important question is how close the actual ANOS is to the ANOS of 16819.9 determined using these estimates. For the actual true parameter values of $p_0 = 0.01$ and $\rho = 0.2$, the true in-control ANOS of this chart is 22629.2. Thus, in this example, there is a significant discrepancy between the actual in-control ANOS and the value computed based on the estimated parameters.

To investigate the effect of errors in estimating process parameters, an additional 99 Phase I data sets were simulated, where each of these data sets had 10,000 in-control observations ($p_0 = 0.01$ and
FIGURE 5. Values of $(\hat{p}_0, \hat{\rho})$ for 100 Phase I Data Sets of Size 10,000.

For each of these data sets, $\hat{p}_0$ and $\hat{\rho}$ were obtained and used to design an approximate MBCUSUM chart with in-control ANOS as close as possible to 16,900. Figure 5 is a plot of the 100 values of $(\hat{p}_0, \hat{\rho})$. We see that many of these values of $(\hat{p}_0, \hat{\rho})$ differ significantly from the true value of $(p_0, \rho) = (0.01, 0.2)$.

The true in-control ANOS values of the 100 MBCUSUM charts designed based on the Phase I data sets were also computed and plotted as a histogram in Figure 6. From Figure 6, we see that many of the true in-control ANOS values are far from the desired value of 16,900.

The practical implication of this is that if you design an MBCUSUM chart based on a Phase I data set of 10,000 in-control observations, then you cannot be confident that the true in-control ANOS will be close to the value that you compute based on the values of the estimated process parameters. This implies, of course, that you need to use a larger Phase I data set to obtain more precise process parameter estimates.

To illustrate the effect of the size of the Phase I data set, the process of generating 100 Phase I data sets, estimating process parameters, and constructing the MBCUSUM chart was repeated using Phase I data sets of size 100,000. Figure 7 is a plot of the 100 values of $(\hat{p}_0, \hat{\rho})$, and Figure 8 is the histogram of the true in-control ANOS values. We see that, using a Phase I data set of size 100,000 should give process parameter estimates that are reasonably close to the true values and also give a true in-control ANOS value that is reasonably close to the desired value of 16,900. Notice that the plots are scaled exactly the same to make them directly comparable (Figure 7 vs. Figure 5, and Figure 8 vs. Figure 6).

In general, the size of the Phase I data set required to give good results will depend on the values of $p_0$ and $\rho$. In particular, smaller values of $p_0$ will require
a larger Phase I data set, while a smaller Phase I data set will be satisfactory for larger values of $p_0$. However, the results given here for $p_0 = .01$ show that the Phase I data set needs to be quite large, in fact, larger than what may commonly be available in practice. In process monitoring, it is generally true that the required size of the Phase I data set is much larger than what is typically used in practice, so the situation here in the case of monitoring $p$ is similar to the general situation. See Jensen et al. (2006) for a review of the literature on the effect of estimated process parameters on the properties of control charts.

Although it is desirable to have a large Phase I data set, the question arises as to what to do when Phase II monitoring must be started based on a relatively small Phase I data set. Our recommendation in this case is to use the control chart with some caution until more data has been obtained and better estimates of process parameters can be computed. For example, if Phase II monitoring is started based on 10,000 Phase I observations, then the process of estimating parameters and designing the chart could be repeated after, say, another 10,000 observations have been obtained (assuming that all observations appeared to be from an in-control process). This process could be repeated again as additional observations are obtained.

This discussion of the effect of estimated process parameters has been done in the context of the MBCUSUM chart, but similar issues arise for the other charts. The SCRF chart is based on the same model as the MBCUSUM chart, so parameter estimates are also needed for this chart. If the traditional charts are going to be used in the presence of autocorrelation, then it is still necessary to estimate both $p_0$ and $p$ to know how to adjust the limits of these charts to get an acceptable value for the in-control ANOS.

Conclusions and Discussion

This paper has considered the situation in which the binary observations from a process follow a two-state Markov chain model and has shown that the resulting autocorrelation has a deleterious effect on traditional control charts designed for independent observations. In particular, positive autocorrelation leads to many more false alarms than would be expected for independent observations. Thus, when developing control charts to monitor a process with binary observations, it is important to explicitly account for autocorrelation when it is present.

The MBCUSUM chart proposed here is based on a log-likelihood-ratio statistic derived from the two-state Markov chain model. The MBCUSUM chart can be well approximated using a Markov chain, and this allows the MBCUSUM chart to be set up to have specific desired statistical properties.

The numerical results given here showed the overall superior performance of the MBCUSUM chart over traditional charts, which ignore any autocorrelation in the observations, as well as over another control chart that accounts for autocorrelation. An example based on simulated data was used to show how the MBCUSUM chart could be set up using estimates of the process parameters obtained in Phase I. An investigation of the effect of the size of the Phase I data set showed that, as is usual in process monitoring, a large Phase I data set is required for the actual control-chart properties to be close to the properties computed assuming that the estimated parameters are the true parameters.

When monitoring a process with binary observations that can be modeled using the two-state Markov chain model, we recommend using the MBCUSUM chart because it is the most effective control chart for detecting increases in $p$.

Appendix

The properties of the approximate MBCUSUM chart were evaluated by modeling the statistic $C_k^*$ as a Markov chain. Note that all nonpositive values of $C_k^*$ can be grouped together to correspond to one state and that the largest possible value of $C_k^*$ that does not produce a signal is $hc_0 - (1/m)$. By design, the values of $C_k^*$ are multiples of $1/m$, so the number of possible values of $C_k^*$ that need to be considered in determining the transient states is $H = m(hc_0 - (1/m)) + 1 = mh_c$. Each of these possible values must correspond to two states in the Markov chain because we must know the values for $X_{k-1}$ and $C_{k-1}$ to determine $C_k^*$. The total number of transient states is then $2H$, labeled as states 1, 2, ..., $2H$. Then $(X_k, C_k^*) = (0, t/m)$ corresponds to state $2t + 1$ and $(X_k, C_k^*) = (1, t/m)$ corresponds to state $2t + 2$.

If we let $N = (N_1, N_2, \ldots, N_{2H})'$ be the vector of ANOS values corresponding to starting in each of the $2H$ transient states, then $N$ can be obtained in the standard way from $N = (I - Q)^{-1}1$, where $1$ is a column vector of 1’s. If $X_1 = 0$, then $C_1^* = 0$, corresponding to state 1; and if $X_1 = 1$, then $C_1^* =$
because \(X_1 = 0\) with probability \(1 - p\) and \(1\) with probability \(p\). The SSANOS can be calculated from

\[
\pi N, \text{ where } \pi \text{ is the normalized left eigenvector of } Q
\]

that corresponds to the largest eigenvalue (computed for \(p = p_0\)).

The ANOS and SSANOS of the exact MBCUSUM chart could not be obtained exactly by modeling the control statistic \(C_k\) as a Markov chain. We attempted to use a Markov chain approximation in the spirit of Brook and Evans (1972), but found that the accuracy obtained was not as good as that obtained using simulation (for other methods, see Hawkins and Olwell (1998)). Thus, the results given here for the exact MBCUSUM chart are based on simulation with 100 million runs. The out-of-control SSANOS was simulated by generating 10,000 in-control observations for each run and then introducing the increase in \(p\). If a false alarm occurred in a sequence of 10,000 in-control observations, then this sequence was discarded and another sequence was generated.

To model the Bernoulli CUCUM chart as a Markov chain in the presence of autocorrelation, we need two states for each possible value of \(B_k\). The construction of the transition probability matrix is similar to the construction used for the MBCUSUM, only here, the value of \(L_k\) does not depend on \(X_{k-1}\), but the transition probabilities do depend on \(X_{k-1}\).

For modeling the Shewhart chart as a Markov chain, let \(S_{ij} = s\), for \(0 \leq s \leq n\), be the number of defectives observed after observation \(j\) in sample \(i\), for \(j = 1, 2, \ldots, n\). The number of transient states required is \(2n(h_S + 1)\). For a given value of \(S_{ij}\), if the previous observation is a nondefective, then the state is \((i-1)n+j\); otherwise, if the current observation is a defective, then the state is \((i-1)n+j+1\). For the SCRF chart, Shepherd et al. (2007) used a Markov chain with three transient states to obtain the expected number of \(Y\)'s until a signal. They evaluated the SCRF chart using the expected number of \(Y\)'s, and did not consider the ANOS or the steady-state properties of the SCRF chart. To obtain the ANOS of the SCRF chart, we used this Markov chain with three transient states to find the expected number of times each state is occupied, and then obtained the ANOS by using the expected number of observations corresponding to each state (the expected values of \(Y_j \geq 1\)). The SSANOS must account for the fact that the shift in \(p\) can occur anywhere within a sequence of nondefectives, so we used simulation (with 100 million runs) to obtain the SSANOS of the SCRF chart.

References


