Technical Report 08-3, Department of Statistics, Virginia Tech

Monitoring Markov Dependent Observations with a Log-Likelihood Based CUSUM

Shabnam Mousavi[1,2] and Marion R. Reynolds, Jr.[3]

[1]Department of Statistics, *The Pennsylvania State University,* University Park, PA 16802-2111, U.S.A., [2]Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany, [3]Departments of Statistics and Forestry, *Virginia Polytechnic Institute and State University,* Blacksburg, VA 24061-0439, U.S.A.

Abstract

When control charts are used to monitor a proportion p it is traditionally assumed that the binary observations are independent. The work that has been done on monitoring autocorrelated binary observations has assumed a two-state Markov chain model with first-order dependence. We investigate the problem of monitoring p for such observations. We show that the most efficient chart for independent observations, the Bernoulli CUSUM chart, along with the traditional Shewhart chart, are not robust to autocorrelation. One approach to dealing with autocorrelation is to adjust the control limits of the traditional charts, but this does not produce the most efficient charts for detecting changes in p. We develop a more efficient log-likelihood-ratio based CUSUM chart for monitoring binary observations that follow the two-state Markov chain model. We show that this CUSUM chart can be well approximated by using a Markov chain that allows calculation of the properties of this chart. We also show that this CUSUM chart has better overall statistical performance than other charts available in the literature.

Keywords: Autocorrelated Observations, Binary Data, CUSUM, Markov Chain.

Introduction

Control charts have long been used for monitoring industrial processes to detect undesirable changes in these processes. Data on the quality characteristics of a process of interest are usually measured in two forms: continuous measurements (traditionally called *variables* data), and discrete or count data (traditionally called *attributes* data). Count data that take only two values, 0 and 1, can be called *binary* data. Binary observations may arise as the natural outcome of the inspection process. For example, if a component is tested by plugging in to see if it activates, then a binary observation is obtained. In this type of situation the two values for the binary observation are frequently labeled nondefective and defective. In some applications continuous measurements, such as dimensions of a component, are obtained, and the inspected components are classified into the two categories of conforming to the standards (when all dimensions are within specifications), and nonconforming (when at least one dimension is outside of specifications). For convenience we use the terms nondefective and defective for the values 0 and 1, respectively.

The application of process monitoring techniques has now expanded far beyond the traditional industrial setting. For example, Woodall (2006) reviews applications of control charts in health-care monitoring. In health-care monitoring the outcomes can be naturally binary. For example, the results of a certain treatment are cured or not cured. In the marketing arena, the requests received by a customer service department that are/are not answered within the standard reply period, or the deliveries that are/are not sent to the correct address produce binary observation data.

Most published work on control charts, and in particular 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. It has been shown that the quality of items are often serially dependent (see Broadbent (1958)), and that the existence of correlation has an adverse effect on the performance of the monitoring tools (such as control charts) that are designed based on assuming independent data (see Deligonul and Mergen (1987)). Alwan and Roberts (1995) provide a summary of over two hundred quality control applications, in which the data violate the underlying assumptions of control charts, such as independence, which in turn leads to misplaced control limits. Their report highlights the fact that violating assumptions can be traced in a wide variety of practical applications.

Although there are now a number of research papers concerned with autocorrelation in control charts, most of these deal with continuous random variables. The published work that deals with autocorrelated binary data is based on the assumption 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). For this model, Bhat and Lal (1990) showed how to determined the upper and lower control limits of a Shewhart control chart. Their chart is based on the number of defective items in sequential samples taken far enough apart for the samples to be considered independent. For the case of 100% inspection, Blatterman and Champ (1992) evaluated a Shewhart chart based on the number of nondefective items between defective items. Champ, Blatterman, and Rigdon (1994) proposed an attribute CUSUM chart for monitoring the proportion defective, based on the same random variable and determined the run-length distribution of one-sided and two-sided charts. Shepherd, Champ, Ridgon, and Fuller (2007) provide two control charts, both of which plot the number of nondefective items before a defective item. One signals as soon as this value falls outside a certain limit, whereas the other one waits for two out-of-limit values to produce a signal. All the aforementioned work is based on using a two-state Markov chain model for first-order serially dependent binary observations, and then using a sequence of independent random variables in constructing the control statistic. Lai, Xie, and Govindaraju (2000) studied the effect of Markov dependence in a high quality environment on the mean and variance of the number of observations to obtain a defective item. This random variable is one plus that considered by Blatterman and Champ (1992), and reduces to the geometric random variable when correlation does not exist. Lai et al. (2000) illustrate the effect of serial dependence on the lower and upper control limits of a Shewhart chart based on a two-state Markov model.

The main objective of this paper is to develop CUSUM charts for monitoring a process in which the observations are binary and follow a two-state first-order Markov chain model. We consider the situation in which a continuous stream of binary observations is available for process monitoring (as would occur with 100% inspection of all output from the process). It is assumed that these binary observations become available individually, so the CUSUM charts can be based on samples of $n = 1$.

We show that the best control chart for independent observations, the Bernoulli CUSUM chart (see Reynolds and Stoumbos (1999) and (2000)), along with the traditional Shewhart chart based on grouping observations into samples of $n > 1$, are not robust to autocorrelation. Thus there is a need for control charts that explicitly account for autocorrelation.

The CUSUM chart that we propose is based on a statistic which is derived by using the log-likelihood-ratio from the two-state Markov chain. We show that our CUSUM chart, called the MBCUSUM chart, can be well approximated by a CUSUM chart that is itself a Markov chain, thus allowing the MBCUSUM chart to be set up to have specified statistical properties.

We show that the MBCUSUM chart is more efficient than the Bernoulli CUSUM chart and the traditional Shewhart chart, both of which ignore any autocorrelation in the observation. We also show that the MBCUSUM chart is more efficient than a chart recently investigated by Shepherd et al. (2007).

We next define the two-state Markov chain model and define performance measures for control charts. Then we define some control charts that have been used for independent binary observations and investigate their robustness. Finally, we develop the MBCUSUM chart and do performance comparisons with other control charts.

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, which we call nondefective and defective, respectively. We are referring to these observations as binary observations, rather than Bernoulli observations, because "Bernoulli" is usually associated with the case in which the observations are independent.

A two-state Markov chain model has only two states, so the transition probability matrix has four elements, p_{ij}, $i, j = 1, 2$. The rows must sum to one, so this matrix can be characterized using only two parameters. This model has traditionally been parameterized using the parameters $p_{01} = P(X_k = 1 | X_{k-1} = 0)$ and $p_{10} = P(X_k = 0 | X_{k-1} = 1)$ (see, for example, Bhat (1984) or Bhat and Lal (1990)), where p_{01} is labeled *a*, and p_{10} is labeled *b*). For this model the long run proportion defective, $p = P(X_k = 1)$, can be expressed as $p = p_{01} / (p_{01} + p_{10})$, and the correlation coefficient ρ between successive observations can be expressed as $\rho = 1 - (p_{01} + p_{10})$.

For quality control applications, it seems more natural to us to directly parameterize the process in terms of p and ρ, instead of p_{01} and p_{10}. This allows the process to be characterized with the traditional parameter p representing the proportion defective and the parameter ρ representing the level of autocorrelation in the process. Then p_{ij} can be obtained from p and ρ using the expressions

$$p_{00} = 1 - p(1-\rho) \tag{1}$$

$$p_{01} = p(1-\rho) \tag{2}$$

$$p_{10} = (1-p)(1-\rho) \tag{3}$$

$$p_{11} = 1 - (1-p)(1-\rho). \tag{4}$$

Here, of course, $p_{00} + p_{01} = 1$ and $p_{10} + p_{11} = 1$.

To apply this model in practice requires that the in-control values of the parameters in the model be estimated during a Phase I analysis when process data are collected for this purpose. Shepherd et al. (2007) discuss the estimation of the parameters in this model. Here we assume that the Phase I data set is large enough that any error associated with process parameter estimation can be neglected.

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, but this process change does not affect the value of ρ. In some applications detecting a decrease in *p* may also be of interest, but here we do not consider the problem of detecting decreases in *p*.

The first observation X_1 will be observed without knowing the value of a previous observation. Thus, 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.

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. Here we are comparing control charts based on different sample sizes, so different control charts with the same value of the ARL will not necessarily have the same expected number of observations to a signal. Thus, we use the *average number of observations to signal* (ANOS) instead of the ARL. The ANOS is defined as the expected number of observations from the start of process monitoring until a signal by the control chart (see Reynolds and Stoumbos (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* is above p_0 then we want a small ANOS corresponding to fast detection of this out-of-control situation.

If the ANOS computed for some $p > p_0$ is used as a measure of out-of-control performance, then there is the implicit assumption that the increase in p is present when monitoring starts. However, in most applications it is likely that any increase in p will occur some time after monitoring has started. Some control charts, such as CUSUM charts, accumulate information over time, and the control statistics of these charts may not be at their starting values when the increase in p occurs. In this situation a more reasonable representation of the expected detection time can be obtained using 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 or stationary distribution of this statistic, conditional on no false alarms.

When a control chart is based on samples of $n > 1$ the increase in p may occur in the middle of a sample, so this possibility must be incorporated in the computation of the SSANOS. In particular, we assume that when the increase occurs within a sample of n observations, the position of the shift within this sample has a uniform distribution.

Methods for evaluating the ANOS and SSANOS of the control charts being considered in this paper are discussed in the Appendix.

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), 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 samples of n observations. For example, if n = 100 then $X_1, X_2, \ldots, X_{100}$ would constitute the first sample, $X_{101}, X_{102}, \ldots, X_{200}$ would constitute the second sample, and so on. If S_i is the number of defectives in the i^{th} sample, 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 $S_i \geq h$, for some constant h. In many applications of the Shewhart p chart, the chart parameter h would be determined based on "three-sigma" control limits, but in this paper we choose h 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 (see Bhat and Lal (1990) for the distribution). We evaluated the ANOS and SSANOS of this Shewhart p chart by modeling it as a Markov chain (see the Appendix).

Reynolds and Stoumbos (1999) investigated the performance of the Bernoulli CUSUM chart for monitoring p when there is a continuous stream of binary observations and these observations are independent. The Bernoulli CUSUM chart is based on treating each individual observation as a sample of n = 1. The Bernoulli CUSUM control statistic is based on a sum of log-likelihood-ratio statistics for the independent observations. For observation X_k the log-likelihood-ratio statistic is

$$\begin{aligned} L_k &= \ln \frac{f(X_k \mid p_1)}{f(X_k \mid p_0)} \\ &= \ln\left(\frac{p_1^{X_k}(1-p_1)^{(1-X_k)}}{p_0^{X_k}(1-p_0)^{(1-X_k)}} \right) \\ &= X_k \ln\left(\frac{p_1(1-p_0)}{p_0(1-p_1)} \right) + \ln\left(\frac{1-p_1}{1-p_0} \right) \qquad k = 1, 2, \ldots. \end{aligned} \tag{5}$$

where $p_1 > p_0$ is a value of *p* that should be detected quickly. Thus

$$L_k = \begin{cases} \ln\left(\dfrac{1-p_1}{1-p_0}\right) & \text{if } X_k = 0 \\ \ln\left(\dfrac{p_1}{p_0}\right) & \text{if } X_k = 1, \end{cases} \tag{6}$$

and the CUSUM control statistic is

$$B_k = \max\{0, B_{k-1}\} + L_k, \quad k = 1, 2, \ldots, \tag{7}$$

where $B_0 = 0$. Dividing Equation (7) by $\ln((p_1(1-p_0))/(p_0(1-p_1))$ gives a CUSUM control statistic, say B'_k, of the form given in Reynolds and Stoumbos (1999),

$$B'_k = \max\{0, B'_{k-1}\} + (X_k - \gamma_B), \quad k = 1, 2, \ldots, \tag{8}$$

where

$$\gamma_B = -\ln\left(\frac{1-p_1}{1-p_0}\right) / \ln\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right). \tag{9}$$

This chart signals if $B'_k \geq h$.

The chart parameter p_1 (which, for a given value of p_0, determines the value of γ_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 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 whose possible values are integer multiples of $1/m$, and this will allow the Bernoulli CUSUM chart to be modeled exactly as a Markov chain (see the Appendix for more details). Modeling the Bernoulli CUSUM chart as a Markov chain permits the exact computation of the ANOS and SSANOS.

As an alternative to formulating the problem of monitoring p as one of observing Bernoulli observations, an equivalent way to monitor p is to formulate the problem as one of observing the number of nondefectives between defectives. The number of nondefectives between defectives has a geometric distribution, so these geometric observations can be used to construct control charts. 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 the Bernoulli CUSUM chart when the Bernoulli CUSUM chart starts with a headstart. This equivalence implies that is no need to consider the geometric CUSUM chart here separately from the Bernoulli CUSUM chart.

For the case of independent binary observations, Sego et al. (2007) and Joner et al. (2007) recently 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 that the Bernoulli CUSUM chart has better performance than these schemes in almost all cases.

Robustness of Traditional Charts to Autocorrelation

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

Consider first 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 $h = 5$ for this control chart then the in-control ANOS will be 29134.8 when there is no autocorrelation. An in-control ANOS of 29134.8 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 ρ 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.

Table 1 also has in-control ANOS values for the Bernoulli CUSUM chart. The value of h for the Bernoulli CUSUM chart was adjusted so that the in-control ANOS would be very close to the value 29134.8 for the Shewhart chart when $\rho = 0$. This allows for easy comparisons with the Shewhart chart. Column [2] of Table 1 has in-control ANOS values for the Bernoulli CUSUM chart for the case of $p_1 = .025$, and column [3] has values for the case of $p_1 = .040$.

Table 1. In-control ($p = p_0$) ANOS values for the Shewhart chart and the Bernoulli CUSUM chart as a function of ρ when p_0 = .010 and p_1 = .025 or .040.

ρ	p_{01}	$n =$ p_{11}	Shewhart 100 [1]	Bernoulli CUSUM p_1 = .025 1 [2]	Bernoulli CUSUM p_1 = .040 1 [3]
.00	.01	.01	29134.8	29248.6	29050.8
.05	.0095	.0595	16956.9	18464.7	15784.0
.10	.0090	.1090	11200.4	12661.0	9972.2
.15	.0085	.1585	7987.2	9204.0	6914.5
.20	.0080	.2080	6000.4	6988.4	5108.3
.25	.0075	.2575	4682.8	5487.9	3952.4
.30	.0070	.3070	3763.3	4427.0	3168.1
.35	.0065	.3565	3096.7	3651.1	2612.0
.40	.0060	.4060	2599.0	3068.0	2204.1
.45	.0055	.4555	2219.3	2620.4	1897.4
.50	.0050	.5050	1925.4	2271.3	1662.8
		$h =$	5	5.2459	4.0435
		Number of States =	1200	640	372

From Table 1 we see that neither the Shewhart chart nor the Bernoulli CUSUM chart is robust to autocorrelation. The value of ρ does not have to be very far above 0 to produce an in-control ANOS much lower than what would be expected from the case in which the observations are independent. An in-control ANOS that is much lower than expected implies, of course, that false alarms will occur much more frequently than expected.

Consider next the situation in which a Shewhart p chart based on samples of n = 400 is used to monitor a process with $p_0 = .001$. If $h = 3$ for this chart then the in-control ANOS will be 50739.3 when there is no autocorrelation. An in-control ANOS of 50739.3 corresponds to 126.8 samples when n = 400. Table 2 has the same structure as Table 1

Table 2. In-control ($p = p_0$) ANOS values for the Shewhart chart and the Bernoulli CUSUM chart as a function of ρ when p_0 = .001 and p_1 = .004 or .008.

			Shewhart	Bernoulli CUSUM p_1 = .004	Bernoulli CUSUM p_1 = .080
		n =	100	1	1
ρ	p_{01}	p_{11}	[1]	[2]	[3]
.00	.0010	.0010	50739.3	50759.7	50755.4
.05	.0009	.0510	32517.2	33856.7	30794.6
.10	.0009	.1009	23557.2	24648.6	21896.4
.15	.0009	.1509	18293.4	19024.4	16907.5
.20	.0008	.2008	14874.6	15321.5	13749.0
.25	.0008	.2507	12509.5	12753.6	11596.0
.30	.0007	.3007	10804.6	10907.3	10057.4
.35	.0007	.3507	9543.2	9547.6	8925.0
.40	.0006	.4006	8597.9	8533.6	8079.3
.45	.0006	.4506	7890.2	7778.0	7448.7
.50	.0005	.5005	7371.9	7226.1	6990.0
		h =	3	2.8788	2.3468
		Number of States =	3200	2660	1394

and gives in-control ANOS values for the Shewhart chart and Bernoulli CUSUM chart for the current situation, where the Bernoulli CUSUM chart has $p_1 = .004$ or .008. Table 2 gives the same conclusion as Table 1; the traditional charts are not robust to autocorrelation.

When autocorrelation is known to be present, one approach to dealing with this autocorrelation is to adjust the control limits of the traditional charts to try to give more acceptable values for the in-control ANOS. However, we next develop a CUSUM chart specifically for the case of autocorrelation, and show that this chart has better ability to detect shifts in *p* than the traditional charts when there is autocorrelation.

A CUSUM Chart for Autocorrelated Data

To develop a CUSUM chart for binary observations that follow the two-state Markov chain model we need to develop the log-likelihood-ratio statistics for an increase in *p* from p_0 to p_1. The joint density of $X_1, X_2, \ldots, X_k$ can be written as

$$f(X_1, X_2, \ldots, X_k \mid p) = f(X_1 \mid p) \prod_{i=2}^{k} f(X_i \mid X_{i-1}, p), \tag{10}$$

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

$$L_k = \begin{cases} \ln \dfrac{f(X_1 \mid p_1)}{f(X_1 \mid p_0)} & k = 1 \\ \ln \dfrac{f(X_k \mid X_{k-1}, p_1)}{f(X_k \mid X_{k-1}, p_0)} & k = 2, 3, \ldots. \end{cases} \tag{11}$$

The value of L_k for $k=1$ is given by Equation (5). For the case of $k \geq 2$

$$f(X_k \mid X_{k-1}) = p_{00}{}^{(1-X_{k-1})(1-X_k)} \times p_{01}{}^{(1-X_{k-1})X_k} \times p_{10}{}^{X_{k-1}(1-X_k)} \times p_{11}{}^{X_{k-1}X_k} , \tag{12}$$

where $p_{ij} = P\left(X_k = j \mid X_{k+1} = i\right)$, for $i, j \in \{0,1\}$. Forming the log-likelihood-ratio using Equations (1), (2), (3), (4), and (12) gives

$$L_k = (1 - X_{k-1})(1 - X_k)l_{00} + (1 - X_{k-1})X_k l_{01} + X_{k-1}(1 - X_k)l_{10} + X_{k-1}X_k l_{11}, \tag{13}$$

where

$$l_{00} = \ln\frac{1 - p_1(1-\rho)}{1 - p_0(1-\rho)};\ l_{01} = \ln\frac{p_1}{p_0};\ l_{10} = \ln\frac{1 - p_1}{1 - p_0};\ l_{11} = \ln\frac{1-(1 - p_1)(1-\rho)}{1-(1 - p_0)(1-\rho)}. \tag{14}$$

Thus we see that

$$L_k = \begin{cases} l_{00} & \text{if } X_{k-1} = 0 \text{ and } X_k = 0 \\ l_{01} & \text{if } X_{k-1} = 0 \text{ and } X_k = 1 \\ l_{10} & \text{if } X_{k-1} = 1 \text{ and } X_k = 0 \\ l_{11} & \text{if } X_{k-1} = 1 \text{ and } X_k = 1. \end{cases} \tag{15}$$

The CUSUM control statistic then is

$$C_k = \max\{0, C_{k-1}\} + L_k , \quad k = 1, 2, \ldots , \tag{16}$$

where $C_0 = 0$. A signal is given if $C_k \geq h$. Call this chart the Markov Binary CUSUM (MBCUSUM) chart. Note that L_k and C_k given by Equations (15) and (16), respectively, reduce to (6) and (7) when $\rho = 0$. That is, the MBCUSUM reduces to the Bernoulli CUSUM in the absence of correlation.

Recall that properties of the Bernoulli CUSUM 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*. This allows the Bernoulli CUSUM to be modeled as a Markov chain. For the MBCUSUM chart we adopt a similar strategy to get an approximate MBCUSUM chart.

The approximate MBCUSUM chart is obtained by approximating L_k by a random variable whose possible values are integer multiples of a constant. This produces a CUSUM control statistic that is a lattice random variable that can be modeled as a Markov chain. The approach that we used to approximate L_k is to obtain the integer $m = \text{nint}(|1/l_{00}|)$, where $\text{nint}(\cdot)$ indicates the nearest integer value. Then L_k is approximated by a new statistic, say L_k^* whose possible values are integer multiples of 1/*m*. In particular, for $k \geq 2$,

$$L_k^* = \begin{cases} l_{00}^* = \text{nint}(l_{00}\text{m})/\text{m} & \text{if } X_{k-1} = 0 \text{ and } X_k = 0, \\ l_{01}^* = \text{nint}(l_{01}\text{m})/\text{m} & \text{if } X_{k-1} = 0 \text{ and } X_k = 1, \\ l_{10}^* = \text{nint}(l_{10}\text{m})/\text{m} & \text{if } X_{k-1} = 1 \text{ and } X_k = 0, \\ l_{11}^* = \text{nint}(l_{11}\text{m})/\text{m} & \text{if } X_{k-1} = 1 \text{ and } X_k = 1. \end{cases} \quad (17)$$

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

$$C_k^* = \max\{0, C_{k-1}^*\} + L_k^*, \quad k = 1, 2, \ldots, \tag{18}$$

where $C_0^* = 0$. A signal is given if $C_k^* \geq h$, where h now is taken as an integer multiple of $1/m$. Details about modeling this chart as a Markov chain are given in the Appendix.

Comparisons of Charts

To compare the out-of-control performance of different control charts, we would like to choose the control limits so that the in-control ANOS values are the same. However, the discreetness of the distributions only allows us to get in-control ANOS values that are very close. The Shewhart chart has 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 match the in-control ANOS of the Shewhart chart as closely as possible.

Consider first the situation from Table 1 in which $p_0 = .010$ and the Shewhart chart is based on samples of n = 100 with a control limit of h = 5. When $\rho = .05$ the in-control ANOS of this chart is 16956.9 (this corresponds to 169.6 samples of n = 100). The column labeled [1] in Table 3 gives out-of-control SSANOS values for this chart for various values of $p > p_0$. Note that this chart can signal only after a sample of 100 observations has been obtained. However, the SSANOS can actually be below 100 because the increase in p may occur while the sample is being taken, so the time from the increase in p to the end of the sample may be less than 100. Even though the SSANOS may be below 100, column [1] shows that the SSANOS of this chart is relatively high for very high values of p. The reason is, of course, that if a shift in p

occurs early in the sample of 100, a signal cannot be given until the complete sample of 100 observations have been obtained, regardless of how many defectives may have been observed.

It seems clear that the performance of the Shewhart chart could be improved if curtailed sampling was used. With curtailed sampling a signal would be given as soon as

Table 3. ANOS and SSANOS values for Shewhart and CUSUM charts for detecting increases in p when $p_0 = .010$, $p_1 = .025$ or .040, and $\rho = .05$.

			Shewhart		CUSUM, p_1 = .025			CUSUM, p_1 = .040		
			Standard	Curtailed	Bernoulli	Exact MBCUCUM	Approx MBCUCUM	Bernoulli	Exact MBCUCUM	Approx MBCUCUM
		n =	100	100	1	1	1	1	1	1
p	p_{01}	p_{11}	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
.010	.0095	.0595	16956.9	16935.5	16977.5	16869.6	16850.7	17046.1	16908.0	16914.2
.015	.0142	.0643	4226.2	4210.2	2351.4	2195.9	2200.7	3155.0	2853.7	2876.8
.020	.0190	.0690	1678.5	1658.3	848.3	796.8	798.0	1102.0	996.0	1004.6
.025	.0238	.0738	870.6	848.3	473.3	448.0	448.1	559.9	512.3	515.7
.030	.0285	.0785	536.6	512.8	322.3	306.7	306.6	353.0	326.4	327.9
.040	.0380	.0880	281.4	254.9	195.2	187.3	187.1	195.1	182.9	183.1
.050	.0475	.0975	190.3	161.2	139.8	135.2	134.9	133.4	126.1	126.1
.070	.0665	.1165	125.4	91.7	89.1	87.5	87.3	81.4	78.1	77.9
.100	.0950	.1450	95.6	56.6	57.7	57.9	57.7	51.3	50.3	50.1
.200	.1900	.2400	70.5	25.4	26.7	28.4	28.3	23.3	23.8	23.7
.300	.2850	.3350	62.7	16.3	17.5	19.9	19.9	15.3	16.3	16.3
.400	.3800	.4300	58.8	12.0	13.1	16.0	16.0	11.4	13.0	13.0
.500	.4750	.5250	56.5	9.5	10.4	14.0	13.9	9.1	11.1	11.1
.700	.6650	.7150	53.9	6.7	7.4	12.4	12.3	6.4	9.5	9.5
.900	.8550	.9050	52.5	5.2	5.7	13.0	12.9	5.0	9.4	9.4
		h =	5	5	5.1475	4.3058	4.2899	4.1087	5.1496	5.1176
		Number of States =	1200	1200	628	–	592	378	–	348

five defectives have been found in a sample without waiting until the end of the sample. Column [2] in Table 3 gives the in-control ANOS and out-of-control SSANOS values for the Shewhart chart with curtained sampling. We see that using curtailed 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. For example, if p increases to .50, then the SSANOS for the standard sampling method is 56.5, while it is only 9.5 for curtailed sampling. Although it is clear that curtailed sampling can be quite beneficial if there is a large shift in p, curtained sampling has rarely been used in applications.

In Table 3, 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 MBCUSUM chart when $p_1 = .025$. Columns [6], [7], and [8] correspond to the case of $p_1 = .040$. 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 3 we see that the exact MBCUSUM is very slightly better than the approximate MBCUSUM 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 it should be of no practical concern. Thus it appears that the approximate MBCUSUM can be used instead of the exact MBCUSUM with negligible effect on the SSANOS performance.

From Table 3 we see that the MBCUSUM chart has better performance than the Bernoulli CUSUM chart except for very large shifts in p. Very large shifts in p were included here 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 MBCUSUM 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 3 show that the CUSUM charts have much better performance than the Shewhart chart for small shifts in *p*. The CUSUM chart also have much better performance than the Shewhart chart for large shifts in *p* unless curtailed sampling is used in the Shewhart chart. As expected, the performance of the CUSUM charts depends on the choice of p_1, with $p_1 = .025$ giving better performance for small shifts, and $p_1 = .040$ giving better performance for larger shifts. Even when curtailed sampling is used with the Shewhart chart, the Shewhart chart is uniformly worse than the Bernoulli CUSUM chart with $p_1 = .040$.

Next consider the situation in which $p_0 = .010$ and $\rho = .20$. If a Shewhart chart with n = 100 and a control limit of h = 5 is used then, from Table 1, we see that the in-control ANOS of this chart is only 6000.4 (corresponding to 60.0 samples of n = 100). To obtain a larger in-control ANOS consider a control limit of h = 6, which gives an in-control ANOS of 16890.0 (corresponding to 168.9 samples of n = 100). Table 4 gives in-control ANOS and out-of-control SSANOS values for Shewhart and CUSUM charts for this situation, where h for the CUSUM charts has been adjusted to give an in-control ANOS approximately the same as for the Shewhart chart. The structure of Table 4 is the same as for Table 3.

The basic conclusions from Table 4 are similar to those from Table 3. In particular, there is negligible difference between the exact and approximate MBCUSUM charts, the MBCUSUM chart is better than the Bernoulli CUSUM chart except for very large shifts, the CUSUM charts are much better than the Shewhart chart for small shifts, and the

CUSUM charts are much better than the Shewhart chart for large shifts unless curtailed sampling is used with the Shewhart chart.

The in-control ANOS values in Table 4 are close to those of Table 3, so comparisons can be made between the case of $\rho = .05$ in Table 3 and the case of higher correlation

Table 4. ANOS and SSANOS values for Shewhart and CUSUM charts for detecting increases in p when p_0 = .010, p_1 = .025 or .040, and ρ = .20.

			Shewhart		CUSUM, p_1 = .025			CUSUM, p_1 = .040		
			Standard	Curtailed	Bernoulli	Exact MBCUSUM	Approx MBCUSUM	Bernoulli	Exact MBCUSUM	Approx MBCUSUM
		n =	100	100	1	1	1	1	1	1
p	p_{01}	p_{11}	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
.010	.0080	.2080	16890.0	16863.3	16830.1	16821.1	16814.2	16815.9	16943.0	16945.9
.015	.0120	.2120	5509.6	5493.1	2956.0	2394.3	2391.6	3989.9	3057.0	3034.0
.020	.0160	.2160	2476.1	2454.8	1114.6	896.9	896.3	1499.4	1106.9	1098.5
.025	.0200	.2200	1351.7	1328.4	625.5	509.5	509.3	773.1	580.1	576.8
.030	.0240	.2240	842.7	818.2	426.0	350.3	350.3	487.6	373.2	371.7
.040	.0320	.2320	427.2	400.8	257.6	214.8	214.8	268.3	210.8	210.6
.050	.0400	.2400	272.3	244.2	184.2	155.4	155.4	182.8	146.0	146.1
.070	.0560	.2560	160.9	129.3	117.3	100.8	100.8	111.1	90.8	91.0
.100	.0800	.2800	111.9	75.5	76.0	66.9	66.9	70.0	58.6	58.9
.200	.1600	.3600	77.1	32.7	35.2	33.4	33.4	31.2	28.2	28.4
.300	.2400	.4400	67.2	20.9	22.7	24.2	24.1	19.9	20.1	20.2
.400	.3200	.5200	62.3	15.3	16.6	20.2	20.2	14.6	16.7	16.8
.500	.4000	.6000	59.4	12.1	13.1	18.4	18.3	11.6	15.1	15.2
.700	.5600	.7600	56.0	8.5	9.3	18.5	18.3	8.2	14.7	15.1
.900	.7200	.9200	54.2	6.6	7.2	27.5	26.6	6.4	19.7	21.0
		h =	6	6	6.6721	4.1393	4.1585	5.5000	4.9883	5.0488
		Number of States =	1400	1400	814	–	682	506	–	414

($\rho = .20$) in Table 4. We see that detecting a given shift in p is harder when the correlation is higher. For example, the approximate MBCUSUM chart with $p_1 = .040$ in

Table 3 requires an average of 183.1 observations to detect a shift to $p = .040$, while for the higher correlation in Table 4 this chart requires an average of 210.6 observations.

Next consider the situation from Table 2 in which a Shewhart p chart based on samples of n = 400 is used to monitor a process in which $p_0 = .001$. If $h = 3$ for this chart then the in-control ANOS will be 32517.2 when $\rho = .05$ (corresponding to 81.3 samples when n = 400). Table 5 gives in-control ANOS and out-of-control SSANOS values for charts for this situation. The structure of Table 5 is similar to that of Tables 3 and 4.

When $p_0 = .001$ and $\rho = .20$, a Shewhart chart with n = 400 and a control limit of h = 3 has an in-control ANOS of 14874.6 (corresponding to 37.2 samples of n = 400). To obtain a larger in-control ANOS consider a control limit of h = 4, which gives an in-control ANOS of 50369.9 (corresponding to 125.9 samples of n = 400). Table 6 gives in-control ANOS and out-of-control SSANOS values for charts for this situation.

The conclusions from Tables 5 and 6 for the case of $p_0 = .001$ are similar to the conclusions from Tables 3 and 4 for the case of $p_0 = .010$. Thus we conclude that when there is autocorrelation, the MBCUSUM chart gives better overall performance than either the Bernoulli CUSUM chart (which is designed for independent observations) or the traditional Shewhart chart (which is based on grouping observations into samples of n).

A Chart Proposed by Shepherd et al. (2007)

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

Table 5. ANOS and SSANOS values for Shewhart and CUSUM charts for detecting increases in p when p_0 = .001, p_1 = .004 or .008, and $\rho = .05$.

			Shewhart		CUSUM, p_1 = .004			CUSUM, p_1 = .008		
			Standard	Curtailed	Bernoulli	Exact MBCUSUM	Approx MBCUSUM	Bernoulli	Exact MBCUSUM	Approx MBCUSUM
		n =	400	400	1	1	1	1	1	1
p	p_{01}	p_{11}	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
.001	.0009	.0510	32517.2	32389.8	32502.8	32523.2	32575.6	32562.4	32517.3	32528.0
.002	.0019	.0519	7153.0	7057.0	5190.8	4752.6	4758.2	6425.7	5529.2	5525.5
.003	.0029	.0529	3119.1	3004.2	2161.2	1988.9	1990.5	2629.3	2264.2	2262.5
.004	.0038	.0538	1828.2	1704.3	1290.6	1193.8	1194.6	1484.2	1299.1	1298.2
.005	.0047	.0548	1261.7	1131.3	909.2	842.9	843.3	993.9	882.6	882.3
.006	.0057	.0557	963.2	827.3	700.2	649.2	649.5	735.7	661.1	660.9
.007	.0067	.0566	785.9	645.1	569.2	527.3	527.5	580.2	526.6	526.5
.008	.0076	.0576	671.5	526.2	479.7	443.7	443.8	477.7	437.2	437.2
.010	.0095	.0595	536.1	382.8	365.2	336.5	336.5	351.9	327.0	327.0
.015	.0142	.0643	396.3	228.1	228.8	209.3	209.3	211.6	202.8	202.9
.020	.0190	.0690	340.6	163.1	165.6	152.1	152.1	151.5	148.6	148.7
.025	.0238	.0738	310.2	127.1	129.0	119.9	119.9	118.4	117.8	117.9
.030	.0285	.0785	290.6	104.1	105.1	99.3	99.3	97.4	97.8	97.9
.040	.0380	.0880	266.4	76.4	76.2	74.4	74.4	72.2	73.2	73.3
.050	.0475	.0975	251.9	60.3	59.7	59.7	59.7	57.5	58.5	58.6
.060	.0570	.1070	242.3	49.8	49.0	49.9	49.9	47.8	48.8	48.9
.080	.0760	.1260	230.3	37.0	36.2	37.9	37.9	35.8	36.7	36.8
.100	.0950	.1450	223.2	29.4	28.8	30.7	30.7	28.6	29.4	29.5
.200	.1900	.2400	208.9	14.5	14.3	16.5	16.5	14.2	14.9	15.0
.300	.2850	.3350	204.1	9.6	9.6	12.0	12.0	9.5	10.2	10.3
.500	.4750	.5250	200.3	5.7	5.7	9.2	9.2	5.7	6.9	6.9
		h =	3	3	2.8506	3.5733	3.5743	2.3906	4.2466	4.2533
		Number of States =	3200	3200	2634	–	2502	1420	–	1276

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.

Table 6. ANOS and SSANOS values for Shewhart and CUSUM charts for detecting increases in p when p_0 = .001, p_1 = .004 or .008, and ρ = .20.

			Shewhart		CUSUM, p_1 = .004			CUSUM, p_1 = .008		
			Standard	Curtailed	Bernoulli	Exact MBCUSUM	Approx MBCUSUM	Bernoulli	Exact MBCUSUM	Approx MBCUSUM
		n =	400	400	1	1	1	1	1	1
p	p_{01}	p_{11}	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
.001	.0008	.2008	50369.9	50220.1	50366.2	50406.8	50398.4	50464.3	50464.5	50463.0
.002	.0016	.2016	13653.9	13557.5	8260.8	6414.0	6409.7	11489.9	7824.8	7824.2
.003	.0024	.2024	6141.5	6026.2	3318.9	2581.8	2580.4	4603.7	3050.6	3050.2
.004	.0032	.2032	3509.3	3387.0	1939.8	1530.7	1530.4	2504.4	1704.4	1704.2
.005	.0040	.2040	2313.9	2187.3	1350.5	1076.1	1076.1	1631.5	1140.2	1140.2
.006	.0048	.2048	1678.2	1548.2	1032.1	827.7	827.9	1186.0	845.5	845.5
.007	.0056	.2056	1302.0	1169.0	834.5	672.4	672.5	924.4	668.5	668.6
.008	.0064	.2064	1061.5	925.4	700.3	566.0	566.2	755.0	552.0	552.0
.010	.0080	.2080	782.7	640.8	530.0	429.8	430.0	550.8	409.1	409.1
.015	.0120	.2120	509.9	354.4	330.7	267.0	267.2	327.6	249.9	249.9
.020	.0160	.2160	411.6	245.2	241.5	192.5	192.6	232.6	181.6	181.6
.025	.0200	.2200	362.7	188.1	190.8	150.1	150.2	180.0	143.5	143.5
.030	.0240	.2240	333.1	152.9	158.0	123.1	123.2	146.4	119.1	119.1
.040	.0320	.2320	298.1	111.3	117.7	91.0	91.1	106.3	89.3	89.3
.050	.0400	.2400	277.5	87.4	93.8	72.6	72.7	83.4	71.7	71.7
.060	.0480	.2480	263.8	72.0	78.0	60.7	60.7	68.7	60.0	60.0
.080	.0640	.2640	246.8	53.2	58.3	46.0	46.0	51.0	45.5	45.5
.100	.0800	.2800	236.5	42.2	46.6	37.3	37.3	40.7	36.9	36.9
.200	.1600	.3600	216.2	13.7	23.2	20.1	20.1	20.3	19.9	19.9
.300	.2400	.4400	209.4	10.2	15.5	14.7	14.7	13.5	14.5	14.5
.500	.4000	.6000	204.0	8.2	9.3	11.3	11.3	8.1	11.2	11.2
		h =	4	4	4.0931	3.8031	3.8101	3.6128	4.6124	4.6180
		Number of States =	4000	4000	3782	–	3170	2146	–	1644

One control chart proposed by Shepherd et al. (2007) signals if a value of Y_j falls below a lower control limit, and the second control chart signals if two consecutive values of Y_j fall below a lower control limit. 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, here we are using the SSANOS as the measure of out-of-control performance, and 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. Thus we can simplify the specification of these charts by using one control limit h for all Y_j, $j \geq 1$. The second control chart seems to have the best properties, so we consider a control chart that signals if two consecutive values of Y_j fall below h, and, for convenience, refer to this chart as the SCRF chart (for Shepherd, Champs, Ridgon, and Fuller).

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, Woodall, and Reynolds (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.

We now consider the performance of the SCRF chart relative to the performance of the MBCUSUM chart when there is autocorrelation. We were not able to find parameters of the SCRF chart that give a close match to the in-control ANOS values used in Tables 3 – 6, so we chose four SCRF charts and then adjusted *h* for the approximate MBCUSUM chart to closely match the in-control ANOS of the SCRF charts. In particular, we chose $p_0 = .001$, and for $\rho = .05$ we used *h* = 50 and 200, and for $\rho = .20$ we used *h* = 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 7 (details about obtaining these values for the SCRF chart are given in the Appendix). Values of p_{01} and p_{11} are not given in Table 7, but can be obtained from Tables 5 and 6 for given values of *p* and ρ.

The SSANOS values in Table 7 show that the MBCUSUM chart has better performance than the SCRF chart, except for very large shifts. For example, comparing columns [1] and [2] in Table 7 shows that the MBCUSUM has a lower SSANOS for all values of *p* shown except for $p = .500$. Comparing columns [3] and [4] shows that the MBCUSUM has a lower SSANOS except for $p \geq .050$. Now $p = .500$ corresponds to a 500-fold increase in *p* above the in-control value of .001, and $p = .050$ corresponds to a 50-fold increase. It does not seem likely that such large increases in *p* would be of much concern in most applications. We included large values of *p* in the tables just to show how the charts perform in extreme situations. Even in the cases of very large values of *p* where the SCRF chart is better than the MBCUSUM, the SCRF chart is not dramatically better than the MBCUSUM. But for smaller values of *p*, which are presumably of more interest, the MBCUSUM can be dramatically better than the SCRF chart.

Table 7. ANOS and SSANOS values for the SCRF and MBCUSUM charts for detecting increases in p when $p_0 = .001$, $p_1 = .008$, and $\rho = .05$ or $.20$.

	$\rho = .05$				$\rho = .20$			
	SCRF	Approx MBCUSUM	SCRF	Approx MBCUSUM	SCRF	Approx MBCUSUM	SCRF	Approx MBCUSUM
$n =$	–	1	–	1	–	1	–	1
p	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
.001	26641.8	26656.8	124029.0	124620.5	19380.4	19471.3	29494.7	29702.4
.002	5826.6	4975.5	30978.0	12308.2	6894.7	4527.8	13847.9	5728.9
.003	2515.4	2107.0	12867.2	3901.3	3593.9	2082.4	8762.8	2452.3
.004	1445.9	1227.8	6794.1	1979.3	2226.7	1259.7	6249.4	1435.5
.005	968.5	841.1	4133.3	1262.6	1529.5	878.9	4764.7	985.4
.006	712.1	632.9	2762.1	912.6	1125.7	667.3	3790.6	742.2
.007	556.9	505.4	1972.3	710.8	870.4	535.1	3107.2	592.9
.008	454.7	420.3	1480.0	581.1	698.6	445.5	2604.8	493.1
.010	330.9	314.6	927.3	425.4	487.6	332.8	1920.6	368.9
.015	197.3	194.6	417.4	254.7	262.9	202.6	1068.5	227.6
.020	142.0	141.6	248.8	181.4	175.5	144.7	686.7	165.0
.025	111.6	111.3	171.8	140.3	131.1	112.1	480.8	129.4
.030	92.3	91.6	129.5	114.0	104.7	91.1	356.6	106.2
.040	68.8	67.4	86.0	82.7	75.1	66.1	220.4	77.7
.050	54.9	53.0	64.4	65.0	58.9	51.9	151.1	60.9
.060	45.6	43.6	51.6	53.8	48.7	42.8	111.0	49.9
.080	34.2	31.9	37.4	40.4	36.3	31.9	68.8	36.4
.100	27.3	25.1	29.5	32.6	29.0	25.7	48.0	28.6
.200	13.6	12.2	14.6	17.6	14.4	13.6	17.4	14.1
.300	9.1	8.4	9.7	12.8	9.6	9.7	10.5	9.9
.500	5.4	6.1	5.8	9.9	5.8	7.2	6.0	7.3
h	200	4.1133	50	5.7333	100	3.8483	10	4.0843
states	–	1234	–	1720	–	1370	–	1454

Conclusions and Discussion

We have considered the situation in which the binary observations from a process follow a two-state Markov chain model, and have shown that the resulting autocorrelation of the observations has a deleterious effect on standard 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 take autocorrelation into account when it is present.

We have developed a CUSUM chart for the situation in which the autocorrelated binary observations follow a two-state Markov chain model. This CUSUM chart, called the MBCUSUM chart, is derived directly from the log-likelihood-ratio statistics for the first-order Markov dependent binary data. We have shown that this CUSUM chart can be well approximated by a Markov chain by approximating the log-likelihood-ratio statistic by a statistic that is an integer multiple of a constant. This allows for the exact computation of properties such as the ANOS and SSANOS.

Numerical results were given to show that the MBCUSUM chart has better overall performance than the standard Shewhart p chart and the Bernoulli CUSUM chart that were designed for independent observations. Thus, by taking the autocorrelation into account, more effective detection of shifts in p can be obtained.

The performance of the MBCUSUM chart was also compared to the performance of a control chart proposed by Shepherd et al. (2007) for autocorrelated binary observations that follow the two-state Markov chain model. It was found that the MBCUSUM chart has better performance except for very large shifts in p.

When monitoring a process with binary observations that follow the two-state Markov chain model, we recommend using the MBCUSUM chart.

In most applications the primary interest will be in detecting increases in *p*, but in some cases detecting decreases in *p* will also be of interest. In this situation an MBCUSUM chart can be developed for detecting decreases in *p*. Then two MBCUSUM charts, one designed for detecting increases in *p* and the other for decreases in *p*, can be used simultaneously. Two natural extensions of this work would be a general model for autocorrelated binary observations with higher orders of dependence, and control charts for multivariate autocorrelated binary observations.

Appendix

The properties of the approximate MBCUSUM chart were evaluated by modeling the statistic C_k^* as a Markov chain. To see how this works, first note that all non-positive 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 $h-(1/m)$. By design (see Equation (18)), the values of C_k^* are multiples of $1/m$. It follows that the number of possible values of C_k^* that need to be considered in determining the transient states is $H = m\left(h-\left(1/m\right)\right)+1 = mh$. Each of these possible values must correspond to two states in the Markov chain because we must know the value of X_{k-1}, in addition to the value of C_{k-1}^*, to determine the transition probabilities corresponding to the possible values of

C_k^*. The total number of transient states is then $2H$, and we label these states as states $1, 2, \ldots, 2H$. The relationships between the values of (X_k, C_k^*) and the state numbers are

$$(X_k, C_k^*) = (0, (t-1)/m) \iff i = 2t-1$$

$$(X_k, C_k^*) = (1, (t-1)/m) \iff i = 2t$$

for $t = 1, 2, \ldots, H$. The possible transitions are given in Table 8.

As an example to illustrate the construction of the exact and approximate MBCUSUM, suppose that $p_0 = .010$, $p_1 = 0.025$, and $\rho = 0.05$. Then from Equation (15) the possible values for L_k are $l_{00} = -.0145$, $l_{01} = 0.9163$, $l_{10} = -.0153$, and $l_{11} = .2147$. The exact MBCUSUM would be carried out using these values in the CUSUM control statistic given by Equation (16).

Table 8. Transitions and Transition Probabilities for the Approximate MBCUSUM Chart

X_{k-1}	X_k	L_k	Transition	Probability	i
0	0	l_{00}^*	$1 \to 1$ or $i \to i-2$	p_{00}	$3, 5, ..., 2H-1$
0	1	l_{01}^*	$i \to i + 2d_{01} + 1$	p_{01}	$1, 3, ..., 2H - 2ml_{01}^* - 1$
1	0	l_{10}^*	$2 \to 1$ or $i \to i - 2ml_{10}^*$	p_{10}	$2ml_{10}^* + 2, 2ml_{10}^* + 4, ..., 2H$
1	1	l_{11}^*	$i \to i + 2ml_{11}^*$	p_{11}	$2, 4, \ldots, 2H - 2ml_{11}^*$

The approximate MBCUSUM would be constructed using m = int(1/.0145) = int(69.01) = 69, so we approximate L_k with L_k^* whose possible values are integer multiples of 1/69. This gives l_{00}^* = $\text{int}\left((69)(-.0145)\right)/69$ = -1/69, l_{01}^* = $\text{int}\left((69)(.9163)\right)/69$ = 63/69. l_{10}^* = $\text{int}\left((69)(.0153)\right)/69$ = -1/69, and l_{11}^* =

$\text{int}\big((69)(.2147)\big)/69$ = 15/69. For the purpose of illustration, choose $h = 100/69$ (i.e., $H = 100$) so that there are 200 transient states.

The in-control values ($p = p_0$) of conditional probabilities for the binary observations modeled as a two-state Markov chain are

$$p_{00} = 1 - p(1-\rho) = 1 - 0.010(1-0.05) = 0.9905$$

$$p_{01} = 1 - p_{00} = 0.0095$$

$$p_{10} = (1-p)(1-\rho) = (1-0.010)(1-0.05) = .9405$$

$$p_{11} = 1 - p_{10} = 0.0595$$

Each row of the matrix, say $\mathbf{Q}$, of transition probabilities for the transient states has at most two nonzero elements as is shown in Figure 1.

$$\begin{array}{cccc|l}
 & & & j: & 1\;\; 2\;\; 3\;\; \cdots\;\; 31\;\; 32\;\; \cdots\;\; 127\;\; 128\;\; \cdots\;\; 197\;\; 198\;\; 199\;\; 200 \\
 & & & X_k: & 0\;\; 1\;\; 0\;\; \cdots\;\; 0\;\; 1\;\; \cdots\;\; 0\;\; 1\;\; \cdots\;\; 0\;\; 1\;\; 0\;\; 1 \\
 & & & C_k^*: & 0\;\; 0\;\; 1/69\;\; \cdots\;\; 15/69\;\; 15/69\;\; \cdots\;\; 63/69\;\; 63/69\;\; \cdots\;\; 98/69\;\; 98/69\;\; 99/69\;\; 99/69 \\
\underline{t} & \underline{i} & \underline{X_{k-1}} & \underline{C_{k-1}^*} &
\end{array}$$

$$\begin{array}{cccc}
1 & 1 & 0 & 0 \\
1 & 2 & 1 & 0 \\
2 & 3 & 0 & 1/69 \\
2 & 4 & 1 & 1/69 \\
3 & 5 & 0 & 2/69 \\
3 & 6 & 1 & 2/69 \\
\vdots & \vdots & \vdots & \vdots \\
37 & 73 & 0 & 36/69 \\
\vdots & \vdots & \vdots & \vdots \\
85 & 170 & 1 & 84/69 \\
\vdots & \vdots & \vdots & \vdots \\
100 & 199 & 0 & 99/69 \\
100 & 200 & 1 & 99/69
\end{array}
\begin{bmatrix}
p_{00} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & p_{01} & \cdots & 0 & 0 & 0 & 0 \\
p_{10} & 0 & 0 & \cdots & 0 & p_{11} & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
p_{00} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
p_{10} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & p_{00} & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & p_{10} & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & p_{01} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & p_{11} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & p_{00} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & p_{10} & 0 & 0 & 0
\end{bmatrix}$$

Figure 1. The $\mathbf{Q}$ Matrix for the MBCUSUM with $p_0 = 0.010$, $p_1 = 0.025$, $\rho = 0.05$, and $H = 100$.

If we let $\mathbf{N} = (N_1, N_2, \ldots, N_{2H})'$ be the vector of ANOS values corresponding to starting in each of the $2H$ transient states, then $\mathbf{N}$ can be obtained in the standard way from $\mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}$, where $\mathbf{1}$ 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 $C_1^* = l_{01}^*$, corresponding to state $2ml_{01}^* + 2$. Assuming that $C_0^* = 0$ and it is not possible to signal at the first observation ($h_C > l_{01}^*$), the ANOS of interest is

$$\text{ANOS} = 1 + (1 - p)N_1 + pN_{2ml_{01}^*+2},$$

because X_1 is 0 with probability 1 – p and 1 with probability p. The SSANOS can be calculated from $\boldsymbol{\pi}\mathbf{N}$, where $\boldsymbol{\pi}$ is the normalized left eigenvector of $\mathbf{Q}$ 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 Mergen (1978)). 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.

The Bernoulli CUSUM chart with control statistic given by Equation (8) can be modeled as a Markov chain when γ_B given by Equation (9) satisfies $\gamma_B = 1/m$, where m is a positive integer. As in the case of the MBCUSUM, when there is 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. With the Bernoulli CUSUM statistic the value of L_k does not depend on X_{k-1}, but the transition probabilities do depend on X_{k-1}.

The Shewhart chart can also be modeled as a Markov chain. Let $S_{ij} = s$, for $0 \le s \le n$, be the number of defectives observed after observation j in sample i, for $j = 1, 2, \ldots, n$. The number of transient states requited is $2n(h+1)$. 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 $Y_j + 1$, $j \ge 1$). 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

Alwan, L. C., and Roberts, H. V. (1995). "The Problem of Misplaced Control Limits". *Applied Statistics* - *Journal of The Royal Statistical Society Series C* 44, No. 3, pp. 269-278.

Bhat, N. (1984). *Elements of Applied Stochastic Processes, Second Edition*. John Wiley and Sons, New York.

Bhat, N., and Lal, R. (1990). "Attribute Control Charts for Markov Dependent Production Processes". *IIE Transactions* 22, pp. 181-188.

Blatterman, D. K., and Champ, C.W. (1992). "A Shewhart Control Chart Under 100% Inspection for Markov Dependent Attribute Data". *Proceedings of the 23rd Annual Modeling and Simulation Conference*, Pittsburg, PA, pp. 1769-1774.

Broadbent, S. R. (1958). "The Inspection of a Markov Process". *Journal of the Royal Statistical Society Series B-Statistical Methodology* 20, pp. 111-119.

Brook, D., and Evans, D. A. (1972). "An Approach to the Probability distribution of CUSUM Run Length". *Biometrika* 59, pp. 539-549.

Bourke, P. D. (1991). "Detecting a Shift in Fraction Nonconforming Using Run-Length Control Charts with 100% Inspection". *Journal of Quality Technology* 23, pp. 225-238.

Champ, C. W.; Blatterman, D. K., and Rigdon, S. E. (1994). "A CUSUM Quality Control Chart Under 100% Inspection for Markov Dependent Attribute Data". *Proceedings of the Third Industrial Engineering Research Conference*, Atlanta, GA. pp. 34-39.

Chen, R. (1978). " A Surveillance System for Congential Malformations". *Journal of the American Statistical Society*. 73 (362), pp. 323-327.

Deligonul, Z.S., and Mergen, A.E. (1987). "Dependence Bias in Conventional p Charts and its correction with an Approximate Lot Quality Distribution". *Journal of Applied Statistics* 14, pp. 75-81.

Hawkins, D. M. and Olwell, D. H. (1998). *Cumulative Sum Charts and Charting for Quality Improvement*. Springer-Verlag. New York, NY.

Joner, M. D., Woodall, W. H., and Reynolds, M. R., Jr. (2008). "Detecting a Rate Increase Using a Bernoulli Scan Statistic". *Statistics in Medicine*, 27, 2555-2573.

Lai, C. D., Xie, M., and Govindaraju, K. (May 2000). "Study of a Markov Model for a High-Quality Dependent Process". *Journal of Applied Statistics* 27, pp. 461-473.

Reynolds, M. R., Jr., and Stoumbos, Z. G. (1999). "A CUSUM Chart for Monitoring a Proportion When Inspecting Continuously". *Journal of Quality Technology* 31, pp. 87-108.

Reynolds, M. R., Jr., and Stoumbos, Z. G. (2000), "A General Approach to Modeling CUSUM Charts for a Proportion," *IIE Transactions* 32, 6, pp. 515-535.

Reynolds, M. R., Jr., and Stoumbos, Z. G. (2001). "Monitoring a Proportion Using CUSUM and SPRT Control Charts," in *Frontiers in Statistical Quality Control* 6, eds. H.-J. Lenz and P.-Th. Wilrich, Heidelberg, Germany: Springer-Verlag, pp. 155-175.

Sego, L. H., Woodall, W. H., and Reynolds, M. R., Jr. (2007). "A Comparison of Surveillance Methods for Small Incidence Rates". *Statistics in Medicine*, 27, 1225-1247.

Shepherd, D. K., Champ, C. W., Rigdon, S. E., and Fuller, H. T. (2007). "Attribute Charts for Monitoring a Dependent Process". *Quality and Reliability Engineering International* 23, pp. 341-365.

Woodall, W. H. (1997). "Control Charts Based on Attribute Data: Bibliography and Review". *Journal of Quality Technology* 29, pp. 172-183.

Woodall, W. H. (2006). "The Use of Control Charts in Health-Care and Public-Health Surveillance". *Journal of Quality Technology* 38, pp. 89-104.