GLR Control Charts for Monitoring a Proportion

Wandi HUANG

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Statistics

Marion R. Reynolds, Jr. Chair
Pang Du
Inyoung Kim
William H. Woodall

December 6, 2011
Blacksburg, VA

Keywords: Continuous inspection; CUSUM chart;
Moving window; Shewhart chart; Statistical process control;
Steady state average number of observations to signal; Subgroup

GLR Control Charts for Monitoring a Proportion

Wandi HUANG

Abstract

The generalized likelihood ratio (GLR) control charts are studied for monitoring a process proportion of defective or nonconforming items. The type of process change considered is an abrupt sustained increase in the process proportion, which implies deterioration of the process quality. The objective is to effectively detect a wide range of shift sizes.

For the first part of this research, we assume samples are collected using rational subgrouping with sample size $n>1$, and the binomial GLR statistic is constructed based on a moving window of past sample statistics that follow a binomial distribution. Steady state performance is evaluated for the binomial GLR chart and the other widely used binomial charts. We find that in terms of the overall performance, the binomial GLR chart is at least as good as the other charts. In addition, since it has only two charting parameters that both can be easily obtained based on the approach we propose, less effort is required to design the binomial GLR chart for practical applications.

The second part of this research develops a Bernoulli GLR chart to monitor processes based on the continuous inspection, in which case samples of size $n=1$ are observed. A constant upper bound is imposed on the estimate of the process shift, preventing the corresponding Bernoulli GLR statistic from being undefined. Performance comparisons between the Bernoulli GLR chart and the other charts show that the Bernoulli GLR chart has better overall performance than its competitors, especially for detecting small shifts.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Marion R. Reynolds Jr., for his guidance, support and patience throughout my dissertation research. In addition, I would like to thank Dr. Pang Du, Dr. Inyoung Kim, and Dr. William H. Woodall for their insightful suggestions and serving as the committee members. My thanks also go to all the graduate students in Dr. Reynolds' research group, Yiming Peng, Lei Sun, Ning Wang, Sai Wang, Pei Xiao, and Liaosa Xu. I give my most special thanks to my presents for their unlimited love and support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The objective of process monitoring is to determine if any change has occurred in a process and identify the specific time point that the change took place. Control charts are powerful tools for process monitoring that were originally developed in the 1920s for manufacturing processes. More recently, control charts have been used in a wide variety of other areas such as public health surveillance, computer network monitoring, and environmental management. In this dissertation control charts will be discussed using terminology from the context of manufacturing processes, but the same ideas can be easily applied in other areas.

The quality of a product measures the extent to which this product meets the requirements of the customer or conforms to the specifications given for the product. It is desirable that all the products being manufactured are of high quality. However, since variation is inevitable in every process, it could greatly impact the product's quality. Some variation is natural and inherent in the design of a manufacturing process, which is referred to as common cause variation; other variation may be caused by a fundamental change that is not usually present in the process, which is known as special cause variation. If we can promptly detect and identify the special cause variation, it is possible to eliminate it from the process and therefore improve the process quality. A process without the presence of any special cause variation is regarded as in-control or stable.

As one of the most effective tools in statistical process control, a control chart is often used to distinguish between the two types of variation so that we can determine if the process stayed in-control or went out-of-control. Basically, it is a graphical display of how a process performs over time. The idea behind this approach is quite straightforward: we collect samples of a certain size from an on-going process at regular time intervals, calculate control statistics based on those samples, and then plot the sample control statistics on a chart with pre-specified control limits. If a trend emerges in those samples, or if any sample falls outside the control limits, then we declare the process to be out-of-control or unstable, and take actions to find the cause of

variation and adjust the process accordingly. On the other hand, if all the samples are within the control limits, the process is considered to be in-control and no action is needed.

Control charts can be classified into different categories according to the type of quality characteristic that these charts are designed to monitor. A quality characteristic is a measurable characteristic associated with the product that could be used to evaluate its quality. There are two main categories of control charts: those for variable data that display values resulting from measurements on continuous variables, and those for attribute data that represent values generated from counts or the presence or absence of some quality characteristic.

The variable control charts are most often used to monitor the mean and/or variance of a random process, assuming the underlying distribution is normal. One example of this type of chart is the Shewhart $\bar{X}$-chart, which perhaps is the simplest and most widely used control chart. It was proposed by Walter A. Shewhart (1931), who now is generally credited as being the first to introduce control chart approach. However, since the Shewhart chart is based on only the most recent sample, it is not effective in detecting small shifts in the process parameters. Instead, two alternatives to Shewhart charts that have been proposed later take past observations into account. One alternative is to use the cumulative sum (CUSUM) chart developed by Page (1954), while the other one is to use the exponentially weighted moving average (EWMA) chart introduced by Roberts (1959). The choice between those two charts often depends on the user's personal preference. The EWMA chart is known to be approximately equal in its performance to the CUSUM charts, but the latter has the minimax optimality property: it has the fastest detection on average for given shifts compared to other charts with the same in-control performance, under the condition that all the charts are in their worst positions when process shift occurs (see Moustakides (1986)).

Many applications involve collecting attribute data during process monitoring. When an item from a process is examined and at least one of the quality characteristics does not conform to standard, this item is classified as defective or nonconforming; otherwise, it is classified as non-defective or conforming. Most often, we assume that items from a process are independent, and the long run proportion of nonconforming items in this process is of interest. Conventionally, this proportion is represented by p and is usually considered as a measure of product quality in a process – an increase in p is an indication of process deterioration, while a decrease in p

corresponds to a process quality improvement. The most desirable value for p would be zero, which means every item produced in the process is conforming. However, this value is generally not attainable in practice. Instead, we assume that when the process is stable over time and without the presence of any special or assignable cause of variation, p equals an 'in-control' value p_0. In practical applications, p_0 is usually estimated during Phase I, where we collect and analyze the process data retrospectively and construct the control limits. Once we determined that the process data gathered were from a stable process and were representative of the in-control process performance, Phase II monitoring begins. We then use control charts to monitor the process by comparing the control statistic of each sample or item collected from the process to the control limits we obtained from Phase I. In this dissertation, we will only consider Phase II monitoring by assuming that p_0 is known and any estimation error can be ignored. Our primary objective is to detect any special or assignable cause variation that results in declining the process quality and thus an increase in p from p_0.

Traditionally, a process proportion p is monitored by the Shewhart p-chart or the equivalent np-chart. For a general review of control charts for monitoring p, see Woodall (1997) and Szarka and Woodall (2011). In particular, if the total number of nonconforming items in a sample of size n is T, the Shewhart p-chart is based on plotting the statistic T/n, and the Shewhart np-chart is based on plotting T. Since items collected from the process are aggregated into samples first, the control limits and performance measures of Shewhart charts are determined from the binomial distribution. However, due to the discreteness of the binomial distribution, for some specific false-alarm rate, there may not always be a corresponding control limit. In practice the normal approximation to the binomial distribution is often used. For example, the control limits for the Shewhart p-chart are traditionally set at three standard deviations away from p_0, which are often referred to as the 3σ control limits. The disadvantage of using the 3σ control limits is that when p_0 is close to zero and n is not large, the normal approximation is not very accurate, which can result in a false-alarm rate that is quite different from the desired one. Another disadvantage is that the Shewhart p-chart and np-chart are insensitive to small changes in p.

To achieve faster detection of small and intermediate shifts in p, an alternative is to use a CUSUM chart. When items are aggregated into samples, a binomial CUSUM chart is often used

for monitoring p (see, for example, Gan (1993), and Reynolds and Stoumbos (2000)). Its control statistic is a likelihood ratio based on a binomial random variable T, which is the total number of nonconforming items in the current as well as past samples of size n, and a pre-determined out-of-control value $p_1 > p_0$, which is the shift in p that is supposed to be detected quickly by the chart. In contrast to the Shewhart chart, the binomial CUSUM chart has the tuning parameter p_1 which produces a chart that has quick detection for shifts close to p_1. In addition, this chart is less affected by discreteness of the binomial distribution, so it allows for more flexibility in choosing proper control limits. However, performance of the binomial CUSUM chart is not as satisfactory for those shifts not close to p_1, in which case a control chart that has better overall performance for a wide range of shift sizes might be preferred. Another disadvantage of the binomial CUSUM chart is that in practice, to design such a chart to achieve desired statistical properties is rather complicated, because the value of p_1 must be specified and the control limit must be determined. Tables and figures are provided in some references for practitioners to choose the control limit (for example, Gan (1993) gave five plots of the average number of samples to signal (ANSS) against the chart limit for various values of p_1, with sample size $n = 100$ and $p_0 = 0.01, 0.02, \ldots, 0.05$, respectively), however, there are still too many combinations of p_0, p_1, and n that are not covered but could be encountered in applications.

When products are monitored continuously from a process, another type of CUSUM chart called the Bernoulli CUSUM chart could be employed and is usually considered more effective than the binomial CUSUM chart, see Reynolds and Stoumbos (1999). This is due to the fact that for control charts based on binomial random variables, a decision about the process can only be made after a sample of n items were inspected and the corresponding point plotted. If a large process change occurred early in this sample, it could not be detected until the whole sample is collected, which significantly delays the signal time. On the other hand, if a Bernoulli CUSUM chart is applied to items that are examined successively, there is no need to group them into samples of a certain size, in which case a decision about the process change can be made in a more timely manner, and actions can be taken immediately to adjust the process back to in-control. However, as is the same for binomial CUSUM charts, for practical applications of Bernoulli CUSUM charts one also faces the challenge of choosing appropriate control chart parameters, which is never a trivial job and requires knowledge from practitioners.

In summary, traditionally, the Shewhart chart is known to have good performances in detecting any large process change, but it is rather insensitive to small shifts. The CUSUM chart can be tuned so that certain magnitudes of shifts are detected very quickly, though its performance for other shift sizes might not be as satisfactory. Unfortunately, in reality it is rarely the case that the size of a process change is known in advance. In this sense, a control chart that could be used to detect a wide range of shift sizes is preferred.

One option for obtaining good performance over a wide range of shift sizes is to use two or more control charts together as a combination. For instance, a Shewhart chart may be used together with a CUSUM chart, so that the combined chart can detect both large and small shifts quickly (see, for example, Westgard et al. (1977), Lucas (1982), and Wu et al. (2008)). With the Shewhart chart being a special case of the CUSUM chart, this approach is actually equivalent to applying two CUSUM charts simultaneously. Lorden (1971) extended this idea and investigated the case in which multiple CUSUM charts were used at the same time. Also, see Dragalin (1997), Sparks (2000), Zhao et al. (2005), and Han et al. (2007) for more recent work. One disadvantage of applying this approach is that designing such a combination requires specifications of a large number of control chart parameters. Some work has been done on applying this approach in the case of monitoring parameters of the normal distribution, however, there seems to be no such study of using multiple CUSUM charts for monitoring p.

Adaptive control charts are another option for effectively detecting a wide range of shifts. Based on the process data, an adaptive control chart estimates the size of any shift that may occur, and then uses this information to design the chart so that it will respond quickly to the estimated shift size. Adaptive charts can be designed rather flexibly by having multiple tuning parameters, but this could also be a drawback in that it might take much effort for practitioners to select all these parameters and determine an appropriate control limit. For more about the adaptive control charts, see, for example, Sparks (2000), Capizzi and Masarotto (2003), Reynolds and Stoumbos (2006), Shu and Jiang (2006), Jiang et al. (2008), and Wu et al. (2009). Notice that there seems to be no such study of applying an adaptive control scheme for monitoring p.

The primary objective of this dissertation research is to introduce and investigate properties of another type of control chart that is able to detect a wide range of shifts in p. The proposed chart is based on a generalized likelihood ratio (GLR) statistic calculated from the

binomial or Bernoulli distribution. Information such as estimates of the process change-point and shift size that might be useful for post-signal diagnosis are also provided in the process of calculating the GLR statistic. GLR chart has been developed for other situations, such as monitoring the mean and/or variance of normal processes. See, for example, Willsky and Jones (1976), Basseville and Nikiforov (1993), Siegmund and Venkatraman (1995), Lai (1995, 1998, 2001), Apley and Shi (1999), Gombay (2000), Capizzi (2001), Nikiforov (2001), Hawkins et al. (2003), Runger and Testik (2003), Hawkins and Zamba (2005), and Reynolds and Lou (2010) for more details.

The first part of this dissertation investigates the performance of the binomial GLR chart in detecting a wide range of shifts in p relative to that of the Shewhart np-chart, the binomial CUSUM chart, the combined Shewhart-CUSUM chart, as well as combinations of multiple CUSUM charts. It is found that the binomial GLR chart is at least as effective as these other options. A major advantage of the binomial GLR chart over the binomial CUSUM chart lies in that it is no longer necessary to specify the out-of-control value p_1 in advance; instead, the GLR statistic is obtained by maximizing the likelihood over all possible values of $p_1 > p_0$. Because of this, the Binomial GLR chart is designed to have better over-all performance for a wide range of shifts in p than the binomial CUSUM chart, which is tuned to detect a specific out-of-control shift to p_1.

In the second part of this dissertation we consider the situation of continuous inspection where Bernoulli observations are being monitored. A Bernoulli GLR chart with an extra parameter – an upper bound imposed on the estimated shift size $\hat{p}_1$ – was developed and its performance in detecting increases in p is compared to the Shewhart-type CCC-r chart, the Bernoulli CUSUM chart, and the Bernoulli EWMA chart. It is found that the ability of the Bernoulli GLR chart to detect a wide range of shift sizes is not as satisfactory as that in the binomial case; however, if the same value is used for the upper bound and p_1, the Bernoulli GLR chart has a much better performance than the Bernoulli CUSUM chart when $p < p_1$, and is only slightly worse when $p \geq p_1$. Since both charts require almost the same amount of effort to be designed appropriately, we would recommend using the Bernoulli GLR chart instead of the Bernoulli CUSUM chart for its excellent performance in detecting small shifts.

The specific goals of this dissertation research include the following:

1. Develop a binomial GLR chart based on various combinations of the in-control value p_0 and sample size n, and compare its performance to that of the Shewhart chart, the individual binomial CUSUM charts with different values of p_1, the combined Shewhart-CUSUM chart, and the combination of multiple binomial CUSUM charts.
2. Develop a Bernoulli GLR chart for continuous monitoring without subgrouping based on different in-control values of p, and compare the performance to that of the Shewhart-type charts, the Bernoulli CUSUM charts with different values of p_1, and the Bernoulli EWMA charts.
3. Establish the relationship between the GLR chart and a set of CUSUM charts, for both the binomial and Bernoulli cases.
4. Illustrate how to plot the binomial and Bernoulli GLR charts for practical applications and how to interpret the results.
5. Provide general guidelines on the design of the binomial and Bernoulli GLR charts.

This dissertation is organized as follows. In the next chapter, we provide a literature review of some commonly used control charts for both variables and attributes, and a summary of work that have been done previously for the GLR chart. Also included is background information for process monitoring and the sampling plan, as well as the performance metrics that will be used to compare different control charts. Chapter 3 is devoted to derivation of the binomial GLR statistic, along with an illustration of how to plot the binomial GLR chart. This is followed by Chapter 4 that explains the design of a binomial GLR chart in more detail. Chapter 5 presents results from comparing the performance of the binomial GLR chart to the other charts or combinations of other charts. The relationship between the binomial GLR chart and a set of binomial CUSUM charts is also investigated. Chapter 6 begins with the derivation of the Bernoulli GLR statistic, and then presented some sample plots of the Bernoulli GLR chart. In Chapter 7 we discussed how to design a Bernoulli GLR chart if an upper bound is placed on $\hat{p}_1$. We evaluated the performance of the Bernoulli GLR chart with the upper bound relative to the other charts in Chapter 8. In Chapter 9, some concluding remarks and final discussion, as well as topics for future research are provided.

Chapter 2

Background on Control Charts

2.1 Process Description and Sampling Plan

To design a control chart for monitoring p in a manufacturing setting, an appropriate sampling plan must be specified, which usually includes the sample size to use and the frequency of sampling. In many applications, it is infeasible to examine all items produced by the process. One solution is to conduct inspections only during some specified time periods. In this situation, samples of size $n > 1$ are collected from inspections performed within the sampling periods. Then a control chart may be constructed based on the sample statistics and assuming that the number of nonconforming items from each sample follows a binomial distribution.

On the other hand, it happens that in some applications a continuous stream of inspected items is always available, which is referred to as "continuous inspection". One typical example of continuous inspection is when all items produced by a process are examined (100% inspection). However, continuous inspection also exists in the situation with less than 100% inspection. For example, suppose 1,000 items are produced per hour but we can only inspect 60 items per hour because it takes one minute to inspect an item. In this case we are still inspecting continuously and getting a stream of inspected items one minute apart. But we are not doing 100% inspection due to the fact that the production rate of the process is higher than our inspection rate.

If continuous inspection is being used, then we need to choose between constructing a control chart based on samples of size $n = 1$, or artificially aggregating items into samples of size $n > 1$. Reynolds and Stoumbos (1999) considered the effect of using artificial subgroups with CUSUM charts, where they developed a Bernoulli CUSUM chart for $n = 1$, and used a binomial CUSUM chart for $n > 1$. Their study showed that the Bernoulli CUSUM chart was generally better in this case than the traditional approach that applies either a Shewhart np-chart or a binomial CUSUM chart to the sample statistics obtained from subgrouping – neither of these two traditional charts could signal immediately if a large shift occurs in the process.

Note that it is not always straightforward to determine which sampling plan to use – they tend to be highly dependent on the specific processes being monitored. In some cases the sampling plan may have been pre-determined even before a decision was made to monitor the process, so it is not under the control of the people doing the monitoring.

Also notice that for inspection of a continuous stream of items with artificial grouping into samples of size $n > 1$, it is highly likely that the shift in p will occur within a sample. In the case in which inspections are done only at specific time intervals, it is more likely that the shift in p will occur between samples. Since artificial aggregation of continuous inspection is generally not recommended, in this dissertation, for the case of $n > 1$, we will only consider the situation that samples are collected during certain inspection periods, and apply a binomial GLR control chart to the sample statistics. And we assume that shifts in a process always happen between samples, so it would be natural to plot a point on the control chart after each sample. It turns out that conclusions about control chart performance obtained from both situations are similar, although derivations and numerical results are not presented here for the case when there is a continuous stream of items grouped into samples of size $n > 1$. For the case of $n = 1$ dealing with Bernoulli observations, we find that modifications of what has been done for the case of $n > 1$ are required, so the case of $n = 1$ will be studied more comprehensively in the later part of this dissertation.

To define the notation to be used in the dissertation, suppose that samples of size n are collected from a process at regular time intervals with $n > 1$ for the binomial case, and continuous inspection without grouping is implemented with $n = 1$ for the Bernoulli case. This results in a series of random variables T_1, T_2, ..., where T_i is the number of nonconforming items found in sample i. For the binomial case, assume that T_i's are independent and follow a binomial distribution with parameters p_0 and n when the process is in-control up to sample τ, i.e. $T_i \sim$ binomial (n, p_0) for $i = 1, 2, \dots, \tau$, where τ is the process change-point. If there is a shift in p between samples τ and $\tau + 1$, the process is then considered to be out-of-control, and now the T_i's follow a binomial distribution with a shifted parameter, say, p_1, i.e. $T_i \sim$ binomial (n, p_1) for $i = \tau + 1, \tau + 2, \dots$. The Bernoulli process is a special case of the binomial sequence with T_i following a Bernoulli distribution with parameter p_0 when the process is in-control and with parameter p_1 when shift occurs and the process becomes out-of-control. After sample k we

observe and record $T_1,\ T_2, \ldots, T_k$. The goal is to determine if there has been an increase in p before the current sample k.

2.2 Control Chart Performance Measures

The effectiveness of a control chart is measured by the requirement that it should signal quickly if the process is out-of-control, but give very few false-alarms when the process stays in-control. One way to describe the performance of control charts for monitoring a proportion is through the ANSS, which is defined as the expected number of samples until a signal is given since the occurrence of the shift. In addition, we define the average number of observations to signal (ANOS) to be the expected number of individual observations taken until a signal is given since the occurrence of the shift. The ANOS and ANSS are directly related if the sample size is a constant at all sampling points, then $\text{ANOS} = n \times \text{ANSS}$. In this dissertation when we compare control charts with different sample sizes, we use the ANOS since it takes the effect of sample size into account. For other situations like designing a binomial GLR chart with a fixed sample size, the ANSS is used to evaluate various control chart parameters.

When there is no shift in p, both ANSS and ANOS can be used to evaluate the in-control performance of the control charts – larger values of ANSS and ANOS are preferred since they indicate fewer false-alarms. In applications, the control limit of a particular control chart is often chosen so that a desired in-control ANSS(ANOS) value can be achieved.

When evaluating the out-of-control performance of control charts the ANSS(ANOS) can be computed from either the initial state or the steady state. The initial state ANSS(ANOS) is based on the assumption that shifts in the parameter occur at the start of monitoring in Phase II, while the steady state ANSS(ANOS) assumes that the process operates in control for some period of time and then shifts to an out-of-control state after the control statistics have reached steady state. We denote the steady state ANSS and ANOS by SSANSS and SSANOS, respectively. In this dissertation, we will compare the performance of various control charts based on the steady state assumption because in most practical situations, it is more likely that a process would initially operate in control and then change when a shift in p occurs sometime in

the future. The general goal is to minimize SSANSS(SSANOS) values over a wide range of process shifts while maintaining a specific in-control performance.

Simulation was used to evaluate the performance of all the control charts, and the results are based on 1,000,000 iterations to ensure that simulation errors are negligible. We also found that the change-point τ hardly has any effect on the chart performance, as long as τ is reasonably large. Thus, in all the simulations we set the values of τ to be $\tau \geq 1/3\times$(in-control ANSS) (for example, $\tau = 100$ when in-control ANSS = 300), which is sufficiently large for allowing us to ignore its effect on the chart performance.

2.3 General Review of Commonly Used Control Charts

2.3.1 Shewhart Control Charts

The modern theory of control charts was originally proposed by Walter A. Shewhart in 1931, and the control charts developed based on these principles are often called Shewhart control charts. Shewhart charts are very simple to apply, which made them the most widely used control charts. All of the Shewhart charts can be built from a general model. Each point plotted on the chart represents a value for the sample control statistic that measures some quality characteristic taken from the process at each sampling point. The mean and standard derivation of this statistic are computed. A center line is drawn at the value of the mean of this statistic, whereas the upper and lower control limits are drawn typically three standard derivations away from the center line, which indicate the thresholds at which process outcome is statistically unlikely to occur when the process is in control. A signal is given when the control statistic falls outside the control limits, and the process is considered to be out-of-control.

If the control statistic used in Shewhart chart is the sample mean, the corresponding chart is called the Shewhart $\bar{X}$-chart, which is often used to monitor the mean of a normal process. The Shewhart S-chart is based on the sample standard deviation, and is used for monitoring the process variance. In some cases each sample consists of only one observation, i.e. sample size $n = 1$, the Shewhart S-chart cannot be directly applied. The Shewhart moving range chart is

specifically designed for monitoring such a situation, and it is based on the range of two successive observations.

A two-sided Shewhart p-chart, which is used to monitor both increases and decreases in the process proportion p, is constructed based on plotting T_k/n on the chart against the corresponding sample number, with the upper and lower control limits set at

$$h_{\mathrm{S,2-sided}} = p_0 \pm 3\sqrt{\frac{p_0(1-p_0)}{n}}.$$

A signal is given when T_k/n is greater than or equal to the upper control limit or less than or equal to the lower control limit. Notice that the lower control limit for the Shewhart p-chart could sometimes be negative unless a large sample size is used, turning a two-sided chart into a one-sided one with only an upper control limit. Particularly, to monitor increases in p, only the upper control limit h_{S} is needed:

$$h_{\mathrm{S}} = p_0 + 3\sqrt{\frac{p_0(1-p_0)}{n}}.$$

Traditionally people tend to choose h_{S} corresponding to a 3σ limit, which gives an in-control ANSS of 740.8 for normal observations in the one-sided case. But this normal approximation to the binomial distribution may not be very accurate when p_0 is close to zero and n is not very large, in which case the actual in-control ANSS may be far from 740.8. A better alternative is to use the exact binomial distribution to determine h_{S}. However, looking at the binomial probabilities shows that there is a very limited number of possible ANSS values that can be achieved. Thus even with the exact binomial distribution it might not be possible to choose h_{S} to get close to a desired in-control ANSS value. Because there are so few possible in-control ANSS values for the Shewhart chart, we chose one of the ANSS values (and the corresponding value of h_{S}) that seemed to correspond to a reasonable false-alarm rate and used this as a "benchmark". Then, in comparing other control charts to the Shewhart chart, we chose the control limits of the other charts so that the in-control ANSS values of the other charts are close to the in-control ANSS of the Shewhart chart. Choosing control limits that give the charts approximately the same in-control ANSS is required to ensure fair comparisons. In applications,

of course, the control limit of a chart such as the GLR chart would be chosen to closely achieve a desired in-control ANSS without reference to the Shewhart chart.

If instead of T_k/n, we are plotting T_k directly, the corresponding control limit would be h'_S, where $h'_S = n \times h_S$ and the chart signals if $T_k \geq h'_S$. In our discussion later, we will be choosing an integer control limit h'_S for the Shewhart chart that corresponds to a specific in-control performance. For example, $h'_S = 5$ is used for the case when $p_0 = 0.01$ and $n = 100$ and gives an in-control false-alarm rate of 0.0034 per sample.

An important feature of Shewhart charts is that the control statistic only uses information contained in the current sample, but ignores all the past samples. Because of this, Shewhart charts are known to have quick detection for large shifts in the parameter being monitored, but are relatively insensitive to small or intermediate shifts.

2.3.2 CUSUM Control Charts

CUSUM control charts provide an alternative to Shewhart charts when detection of small shifts is of interest. By plotting the cumulative sums of likelihood ratio statistics, CUSUM charts directly incorporate information given by past samples. This idea was first introduced by Page (1954). See Page (1961), Ewan (1963), Brook and Evans (1972), Lucas (1976), Hawkins (1981), Gan (1991), Woodall and Adams (1993), Hawkins and Olwell (1998), and Reynolds and Stoumbos (2000) for more detail. Because it combines information from more than one sample, it is more effective than the Shewhart chart for detecting shifts of small magnitudes. In addition, the CUSUM chart works properly in the situation when there is no subgrouping and the sample size equals one, whereas a much larger sample size is required for the Shewhart chart to perform well.

The most frequently used CUSUM chart is designed for monitoring the mean of a normal process. It is often a two-sided chart based on two separate one-sided control statistics: one is an upper control statistic designed to detect increases in the mean, while the other one is a lower control statistic to detect decreases in the mean. The CUSUM chart signals when any control statistic falls beyond the pre-determined control limits. In fact, the CUSUM chart for monitoring

the process mean is equivalent to applying a sequence of sequential probability ratio tests (SPRTs) for testing $H_0: \mu = \mu_0$ against $H_1: \mu = \mu_1$, where μ_1 serves as a tuning parameter which corresponds to a pre-specified shift size that should be detected quickly by the CUSUM chart.

We may apply CUSUM charts to monitor other types of process variables, such as the process proportion of nonconforming, which can be constructed in the same manner as the CUSUM chart for monitoring the process mean. In particular, we will focus on the CUSUM chart for detecting increases in p, i.e. a one-sided binomial CUSUM chart that is built based on binomial observations (see, for example, Gan (1993)).

The binomial CUSUM chart can be formulated from a sequence of SPRTs for testing $H_0: p = p_0$ against $H_1: p = p_1$. The increment of log likelihood ratio that is accumulated in the binomial CUSUM statistic at sample k is Z_k, which is given by

$$Z_k = \ln \frac{f(T_k|p_1)}{f(T_k|p_0)} = \ln \frac{{p_1}^{T_k}(1-p_1)^{n-T_k}}{{p_0}^{T_k}(1-p_0)^{n-T_k}} = r_2 T_k - n r_1,$$

where

$$r_1 = -\ln \frac{1-p_1}{1-p_0} > 0, \qquad r_2 = \ln \frac{p_1(1-p_0)}{p_0(1-p_1)} > 0, \tag{2.1}$$

and p_1 is the tuning parameter that specifies a particular shift size for which the binomial CUSUM chart is supposed to have fast detection. The CUSUM chart signals when the difference between the cumulative increment of the current sample and that of the previous minimum exceeds the control limit. If we denote this difference in cumulative increments by $C_{\mathrm{Bin},k}$ and the control limit of a binomial CUSUM chart by h_{BinC}, then a one-sided binomial CUSUM chart for detecting increases in p signals if $C_{\mathrm{Bin},k} > h_{\mathrm{BinC}}$, where

$$C_{\mathrm{Bin},k} = \sum_{i=1}^{k} Z_k - \min_{0 \le i < k} \sum_{j=1}^{k} Z_j.$$

Sometimes this CUSUM statistic is written in an iterative form

$$C_{\mathrm{Bin},k} = \max\{0, C_{k-1}\} + Z_k = \max\{0, C_{k-1}\} + (r_2 T_k - n r_1),$$

so that it can be expressed by the previous accumulation $C_{\mathrm{Bin},k-1}$ and the current increment Z_k.

In practice the binomial CUSUM statistic is usually expressed in a different form by dividing both sides of the previous equation by the constant r_2, which leads to a modified CUSUM statistic $C'_{\text{Bin},k}$:

$$C'_{\text{Bin},k} = \max\{0, C'_{\text{Bin},k-1}\} + \frac{Z_k}{r_2} = \max\{0, C'_{\text{Bin},k-1}\} + (T_k - n\gamma),$$

where

$$\gamma = \frac{r_1}{r_2}$$

is usually called the reference value of the CUSUM statistic. In this case, a signal is given if $C'_{\text{Bin},k} > h'_{\text{BinC}}$, where $h'_{\text{BinC}} = h_{\text{BinC}}/r_2$. In this dissertation, in order to illustrate the relationship between the CUSUM and GLR statistics, we will use the CUSUM statistic in the form of $C_{\text{Bin},k}$ and the control limit h_{BinC}.

Notice that the binomial CUSUM statistic $C_{\text{Bin},k}$ can be expressed in the same form as the binomial GLR statistic given in Equation (3.2) in Chapter 3, except that p_1, which serves as the tuning parameter of the CUSUM chart, is pre-specified in the CUSUM statistic instead of estimated as in the GLR statistic:

$$C_{\text{Bin},k} = \ln \frac{\max_{0\le\tau<k} L(\tau, p_1 | T_1, T_2, \ldots, T_k)}{L(\infty, p_0 | T_1, T_2, \ldots, T_k)} = \ln \frac{\max_{0\le\tau<k} \prod_{i=\tau+1}^{k} \binom{n}{T_i} p_1^{T_i} (1-p_1)^{n-T_i}}{\prod_{i=\tau+1}^{k} \binom{n}{T_i} p_0^{T_i} (1-p_0)^{n-T_i}}. \quad (2.2)$$

2.3.3 EWMA Control Charts

Another control chart that considers information contained in the past observations is the EWMA chart, which was first proposed by Roberts (1959). Since then, the properties and performance of EWMA charts have been studied extensively (see, for example, Box et al. (1974), Crowder (1987), Lucas and Saccucci (1990), Crowder and Hamilton (1992), and Yashchin (1993)). Unlike the Shewhart chart, the EWMA chart is very effective in detecting small process shifts.

The EMWA chart is usually used for detecting changes in the process mean. With some minor modification, it could also be used to monitor the process proportion. As is shown in Yeh

et al. (2008), if T_k is the total number of nonconforming items in sample k of size n, the corresponding control statistic E_k is given by

$$E_{\mathrm{Bin},k} = \lambda T_k + (1-\lambda)E_{\mathrm{Bin},k-1}, \quad k = 1,2, \ldots,$$

where λ is a smoothing or tuning parameter satisfying $0 < \lambda \leq 1$. A starting value $E_{\mathrm{Bin},0}$ is needed to compute the first statistic at $k = 1$, and its value is usually specified as the in-control expected value of T_k, i.e.

$$E_{\mathrm{Bin},0} = np_0.$$

In general, values of λ between 0.05 and 0.25 are usually recommended for monitoring normal processes. Sometimes the value of λ can be much smaller if the EWMA chart is used to monitor p based on binomial or Bernoulli distributions. A good rule of thumb is to use smaller values of λ to detect small shifts, and use a large λ value for detecting large shift sizes. When $\lambda = 1$, the EWMA chart is equivalent to the Shewhart np-chart, which considers only the most recent observation.

For detecting increases in p, since values of T_k increase as p increases, only an upper control limit is needed. The exact geometric upper control limit is given by

$$h_{\mathrm{BinE}} = np_0 + w_{\mathrm{Bin}}\sqrt{\frac{np_0(1-p_0)\lambda}{2-\lambda}(1-(1-\lambda)^{2k})},$$

where $k = 1, 2, \ldots$, and w_{Bin} is a multiplier that is determined so that a desired in-control performance can be achieved. The chart signals if $E_{\mathrm{Bin},k} > h_{\mathrm{BinE}}$, given that the exact limit h_{BinE} is used as the control limit.

Note that the term $(1-(1-\lambda)^{2k})$ in h_{BinE} approaches 1 for large k as $k \to \infty$. Thus, an approximate limit h'_{BinE} given below is often used when the process has been operating for some time period and is in its steady state.

$$h'_{\mathrm{BinE}} = np_0 + w_{\mathrm{Bin}}\sqrt{\frac{np_0(1-p_0)\lambda}{2-\lambda}}.$$

The chart signals if $E_{\mathrm{Bin},k} > h'_{\mathrm{BinE}}$ when the approximate limit h'_{BinE} is used.

Also notice that if the same value of w_{Bin} is used in both h_{BinE} and h'_{BinE}, then as $k \to \infty$, $h_{\text{BinE}} \to h'_{\text{BinE}}$, and for small k we have $h_{\text{BinE}} < h'_{\text{BinE}}$, which would result in slightly different in-control ANOS values. In order to maintain constant in-control ANOS, w_{Bin} needs to be adjusted so that we have $h_{\text{BinE}} < h'_{\text{BinE}}$ for small k and $h_{\text{BinE}} > h'_{\text{BinE}}$ as $k \to \infty$. Based on this, using h_{BinE} is better when the process change-point τ is close to 0, and using h'_{BinE} is slightly better when τ is large. When looking at steady state performance, the results are essentially the same whether h_{BinE} or h'_{BinE} is used. On the other hand, since in applications we do not really know when the process changed and whether $\tau = 0$, $\tau = 100$, or $\tau = 10{,}000$, a choice must be made between using h_{BinE} or h'_{BinE} for all values of k. In this dissertation, all performance evaluations are based on the steady state results, thus we decide to always use the approximate limit h'_{BinE} for the EWMA charts.

Crowder and Hamilton (1992) proposed using a reflecting barrier (also known as "reset") in the EWMA chart designed for monitoring a process standard deviation, so that the distance between its control statistic and the control limit has a fixed maximum value (see Woodall and Mahmoud (2005) for more detailed discussion on the inertial effect caused by not using a reset). This reset scheme can be applied to other EMWA charts as well. In particular, the binomial EWMA statistic with reset becomes

$$E'_{\text{Bin},k} = \max\left(np_0, \lambda T_k + (1-\lambda)E'_{\text{Bin},k-1}\right).$$

The reset scheme improves the performance in steady state since now the control statistic cannot be too far away from the control limit when a shift occurs. However, in initial state, the performance with the reset is worse mainly due the fact that, in general, the EWMA chart without the reset has a narrower control limit than the one with the reset, which makes it easier to signal in the initial state situation. In this dissertation, however, since we choose to use the steady state performance as our performance metric, all EWMA charts considered will be using a reset.

The properties of the EWMA chart are actually quite similar to that of the CUSUM chart in terms of their abilities to detect certain sizes of process change. The choice between using an EWMA chart and a CUSUM chart largely depends on user's personal preference. We would generally recommend using the CUSUM chart for its optimality property, as well as for the easier interpretation of the tuning parameter p_1 in the CUSUM chart compared to λ in the

EWMA chart. Therefore, when evaluating the performance of different control charts for monitoring p in the binomial case, we will be focusing on comparisons among the Shewhart charts, CUSUM charts, and GLR charts, but not including the EWMA charts. However, the EWMA charts are included in the Bernoulli case to illustrate the effect of using extremely small values of λ when the Bernoulli observations are observed, which are quite different from the conventionally used range of 0.05 to 0.25.

2.4 Control Charts for Monitoring Bernoulli Processes

2.4.1 Shewhart-type Charts

As is defined in Section 2.1, when continuous inspection is being used as the sampling scheme, we always observe a continuous stream of inspected items, which could be based on either 100% inspection or less than 100% inspection. To monitor a Bernoulli process with continuous inspection, quite often people would artificially group items into samples of size $n > 1$, so that the standard Shewhart p-chart or np-chart can be applied. However, artificial subgrouping has the disadvantage that any decision regarding the process performance can only be made after the whole sample is inspected. This could delay the signal time substantially if a large shift occurs early in a sample (see Reynolds and Stoumbos (1999)). To overcome this, sometimes people try to use a smaller sample size, but it may result in a value of n that is too small for the p-chart or np-chart to detect any change in p effectively. Thus, artificial subgrouping is generally not recommended for monitoring Bernoulli processes if continuous inspection is being used.

Theoretically, a Shewhart chart could be built based on Bernoulli observations with $n = 1$, but this chart would signal any time when a nonconforming item is observed, i.e. $h'_S = 1$, which is not desirable. As alternatives, Shewhart charts based on the geometric or negative binomial distributions were developed and used instead. This type of chart is designed for continuous inspection and monitors the cumulative counts of conforming items produced until a fixed number r of nonconforming items is observed. A point is then plotted on the chart after

observing r nonconforming items. This charting scheme has non-constant lengths for the time intervals between points plotted.

In the case when $r = 1$, the cumulative counts follow a geometric distribution, and Calvin (1983) was the first to propose a Bernoulli chart that was based on geometric observations. Later, Goh (1987) developed a similar chart with $r = 1$ to monitor high yield processes where the number of nonconforming items is measured by parts per million (ppm), which is referred to as the cumulative counts of conforming (CCC) chart and is now widely used to monitor automated processes and processes with near zero-nonconformity. The main advantage of this chart is its simplicity in design and practical application. Calvin (1983) and Goh (1987) used the traditional k-sigma control limits obtained from the geometric distribution. However, since the geometric distribution is not symmetric but highly skewed, Bourke (1991) and Xie et al. (1995) proposed using probability limits instead, which were derived based on the cumulative distribution function (cdf) of the geometric distribution.

Bourke (1991) considered two Shewhart-type charts based on run lengths using a geometric random variable when $r = 1$, and using a negative binomial random variable when $r = 2$. Kaminsky et al. (1992) further investigated the idea of using sums of r ($r > 1$) independent and identically distributed geometric random variables to form negative binomial random variables, and developed a Shewhart chart called the g-chart. The g-chart is similar to the CCC chart but with $r > 1$. Traditional k-sigma control limits calculated from the negative binomial distribution were used for their Shewhart g-chart. Xie and Goh (1997) discussed issues associated with using those k-sigma limits as well as some other technical problems with the Shewhart g-chart, and recommended replacing the k-sigma limits with probability limits obtained from the cdf of the negative binomial distribution. Those probability limits were chosen in the same way as those used for the CCC chart in Xie et al. (1995).

As an extension of the CCC chart, a two-sided CCC chart with $r > 1$ was provided by Xie et al. (1999), and was called the CCC-r chart. This chart is based on the negative binomial distribution, and uses the exact probability limits derived from the cdf of a negative binomial random variable. To select values of r, the authors considered r ranging from 1 to 6. They concluded that to monitor a high yield process with a very small in-control proportion p_0 (e.g.

$p_0 < 0.0001$), a CCC-r with a large r value should not be used. Instead, people should use the CCC-r chart with $r = 1$, which is equivalent to the commonly known CCC chart. For more discussion on using the Shewhart-type CCC or CCC-r charts based on geometric or negative binomial distributions, see, for example, Lucas (1989), Glushkovsky (1994), Chan et al. (1997), Benneyan (2001), Ohta et al. (2001), Wu et al. (2001), Xie et al. (2001), Kuralmani (2002), Xie et al. (2002), Chan et al. (2003), Schwertman (2005), Cheng and Chen (2008), and Lai and Govindaraju (2008).

More details will be given here regarding the design of the Shewhart-type CCC or CCC-r chart. Suppose a process is being examined continuously and the result of inspecting an individual item is immediately available. We record the number of items we observe until a nonconforming item occurs, and denote it by Y, which follows a geometric distribution with cdf

$$F(Y = y, 1, p) = 1 - (1 - p)^y,$$

where $y = 1, 2, \ldots$. Then a CCC chart can be constructed based on Y that plots a point on the chart every time when a nonconforming item is observed. The commonly used probability limits proposed by Bourke (1991) and Xie et al. (1995) for the CCC chart are:

$$\text{LCL} = \frac{\ln(1-\alpha/2)}{\ln(1-p_0)}, \text{and UCL} = \frac{\ln(\alpha/2)}{\ln(1-p_0)},$$

such that $F(\text{LCL}) \approx \alpha/2$ and $F(\text{UCL}) \approx 1 - \alpha/2$, where α is the overall false-alarm rate.

Alternatively, if we record the number of items we observe until r nonconforming items are found, and denote it by W, then $W = \sum_{j=1}^{r} Y_j$ follows a negative binomial distribution with cdf

$$F(W = w, r, p) = \sum_{j=r}^{w} P(Y = j) = \sum_{j=r}^{w} \binom{j-1}{r-1} {p_0}^r (1 - p_0)^{j-r}.$$

In this case a CCC-r chart can be constructed based on W that plots a point on the chart every time r nonconforming items are observed. For a desired false-alarm rate α, the control limits of the CCC-r chart can be obtained numerically as the solutions of:

$$F(\text{LCL}_r, r, p_0) = \sum_{j=r}^{\text{LCL}_r} \binom{j-1}{r-1} {p_0}^r (1 - p_0)^{j-r} = \alpha/2,$$

and

$$F(\mathrm{UCL}_r, r, p_0) = \sum_{j=r}^{\mathrm{UCL}_r} \binom{j-1}{r-1} {p_0}^r (1-p_0)^{j-r} = 1 - \alpha/2.$$

2.4.2 Bernoulli and Geometric CUSUM Charts

CUSUM charts are known to have better performance than Shewhart-type charts in detecting small to moderate shifts in a process. To detect a shift from p_0 to p_1 that occurred in a Bernoulli process, Bourke (1991) used a CUSUM chart based on the geometric distribution. Reynolds and Stoumbos (1999) studied a special case of the binomial CUSUM chart with $n = 1$ and developed a chart that plots a point on the chart after each inspection of an individual item, called the Bernoulli CUSUM chart. For a thorough review of applications of the CUSUM charts to monitor Bernoulli processes, see Szarka and Woodall (2011).

To monitor a process with continuous inspection, the Bernoulli CUSUM chart is constructed based on the individual Bernoulli observations $X_1, X_2, \ldots$, without any kind of subgrouping being used. For detecting increases in p, the upper-sided Bernoulli CUSUM statistic $B_k, k = 1, 2, \ldots$, is given by

$$C_{\mathrm{Bern},k} = \max\{0, C_{\mathrm{Bern},k-1}\} + (r_2 X_k - r_1),$$

where r_1 and r_2 are defined in the same way as those used in the binomial CUSUM statistic $C_{\mathrm{Bern},k}$. The starting value $C_{\mathrm{Bern},0}$ is often set to be 0, but if the fast initial response (FIR) feature is desired to enhance the chart's performance, $C_{\mathrm{Bern},0}$ can take a positive value, which is known as the head start value (see Montgomery (2009), and Szarka and Woodall (2011)). A signal is given if $C_{\mathrm{Bern},k} > h_{\mathrm{BernC}}$, where h_{BernC} is selected to give a desired in-control performance.

In practice, the Bernoulli CUSUM statistic is usually expressed in a different form by dividing both sides of the previous equation by the constant r_2, which leads to a modified CUSUM statistic $C'_{\mathrm{Bern},k}$, given by

$$C'_{\mathrm{Bern},k} = \max\{0, C'_{\mathrm{Bern},k-1}\} + (X_k - \gamma),$$

where

$$\gamma = \frac{r_1}{r_2}$$

is usually called the reference value of the CUSUM statistic. When investigating performance of the Bernoulli CUSUM chart using Markov chains, Reynolds and Stoumbos (1999) found it more convenient to approximate γ by $1/q$, where q is some integer. In this case, a signal is given if $C'_{\text{Bern},k} > h'_{\text{BernC}}$, where $h'_{\text{BernC}} = h_{\text{BernC}}/r_2$. Here, in order to illustrate the relationship between the CUSUM and GLR statistics, we will use the CUSUM statistic in the form of $C_{\text{Bern},k}$ and the control limit h_{BernC}.

Depending on the definition of a geometric random variable, formulation of the geometric CUSUM statistic may take different forms. Following the same definition as used in the previous section, we define $Y_1, Y_2, \ldots$ to be the numbers of conforming items we observe until a nonconforming item is detected in a sequence of Bernoulli observations $X_1, X_2, \ldots$, including the latest nonconforming. For detecting increases in p, after observing the k^{th} nonconforming item, the geometric CUSUM statistic $C_{\text{Geo},k}, k = 1, 2, \ldots$, was given by Chang and Gan (2001) as

$$C_{\text{Geo},k} = \max\{0, C_{\text{Geo},k-1}\} + (-Y_k + \gamma_{\text{G}}),$$

where $C_{\text{Geo},0} = 0, \gamma_{\text{G}} = 1/\gamma = q$. The chart signals if $C_{\text{Geo},k} > h_{\text{GeoC}}$.

Alternatively, if the geometric random variable is defined as the count of items between two nonconforming items but not including the latest nonconforming, which was employed by Bourke (1991), the CUSUM statistic $C'_{\text{Geo},k}$ will have different parameters:

$$C'_{\text{Geo},k} = \max\{0, C'_{\text{Geo},k-1}\} + (-Y_k + \gamma'_{\text{G}}),$$

where $\gamma'_{\text{G}} = \gamma_{\text{G}} - 1$. The chart signals if $C'_{\text{Geo},k} > h'_{\text{GeoC}}$.

Reynolds and Stoumbos (1999), Bourke (2001), Chang and Gan (2001), and Szarka and Woodall (2011) have shown that the Bernoulli CUSUM and geometric CUSUM charts can be set up to be equivalent in detecting increases in p. In particular, if a shift occurs at the beginning of a process, i.e. the initial state performance is considered, a Bernoulli CUSUM chart with a head start being $C_{\text{Bern},0} = (q-1)/q$ is equivalent to a geometric CUSUM chart with $C_{\text{Geo},0} = 0$,

given that the control limits h_{BernC} and h_{GeoC} are appropriately chosen. If a shift occurs after the process is in its steady state, then there is no longer an effect of the head start and a Bernoulli CUSUM chart is equivalent to a geometric CUSUM chart given that a shift is allowed to occur after any item is observaed from the process, which is more realistic than assuming that a shift can only occur after a nonconforming item is observed.

If detecting decreases in p is also of interest, this equivalence no longer holds, and a Bernoulli CUSUM chart is more effective than a geometric CUSUM chart, because the latter can only signal after a nonconforming item is found, while the Bernoulli CUSUM may signal when a long sequence of conforming items is observed.

Notice that Szarka and Woodall (2011) found in their study on the equivalence of the Bernoulli and geometric CUSUM charts that "… the Bernoulli CUSUM chart is more appealing to use than the geometric CUSUM chart in applications involving Bernoulli data". Thus, when the performance of a Bernoulli GLR chart is evaluated later, we will be focusing on comparing its steady state performance to that of a Bernoulli CUSUM chart, given that a geometric CUSUM chart should perform equivalently to a Bernoulli CUSUM chart.

2.4.3 Bernoulli and Geometric EWMA Charts

As an alternative to the CUSUM chart, the EWMA chart is also well known to be more effective than the Shewhart-type charts in detecting shifts of small or intermediate magnitudes. For monitoring Bernoulli processes with a very low level of fraction nonconforming, McCool and Joyner-Motley (1998) first proposed using an EWMA chart based on geometric random variables that were either power or log transformed so that they would be approximately normal. Kotani et al. (2005) studied an EWMA chart based on the negative binomial distribution, using the approximate control limits.

Yeh et al. (2008) derived an EWMA statistic based on Bernoulli observations $X_1, X_2, \ldots$:

$$E_{\text{Bern},k} = \lambda X_k + (1-\lambda) E_{\text{Bern},k-1},$$

where λ is a smoothing or tuning parameter satisfying $0 < \lambda \leq 1$, and $E_{\text{Bern},0}$ is set to be

$$E_{\text{Bern},0} = p_0.$$

For detecting increases in p, the exact and approximate Bernoulli upper control limits are given by

$$h_{\text{BernE}} = p_0 + w_{\text{Bern}}\sqrt{\frac{p_0(1-p_0)\lambda}{2-\lambda}(1-(1-\lambda)^{2k})},$$

and

$$h'_{\text{BernE}} = p_0 + w_{\text{Bern}}\sqrt{\frac{p_0(1-p_0)\lambda}{2-\lambda}},$$

respectively, where $k = 1, 2, ...$, and w_{Bern} is a multiplier that is determined so that a desired in-control performance can be achieved. The chart signals if $E_{\text{Bern},k} < h_{\text{BernE}}$, given that h_{BernE} is used as the control limit, or $E_{\text{Bern},k} < h'_{\text{BernE}}$ if h'_{BernE} is used. Since the value of λ is generally set to be extremely small for monitoring Bernoulli observations, using the exact limit h_{BernE} has a large impact on the performance of an EWMA chart such that its steady state performance could be quite different from the corresponding initial state performance. However, in the approximate limit h'_{BernE} the factor of $(1-(1-\lambda)^{2k})$ is set to be 1, which reduces the effect of using a small λ .The initial state and steady state performance become similar in this case. Therefore, when evaluating the performance of an EWMA chart we will always use the approximate control limit h'_{BernE}.

In another case, let Y be the number of conforming items observed until a nonconforming item occurs, with the last nonconforming included, Y follows a geometric distribution with mean and variance given by

$$E(Y) = 1/p, \text{ and } Var(Y) = (1-p)/p^2.$$

Yeh et al. (2008) showed that an EWMA statistic based on geometric observations can be defined as

$$E_{\text{Geo},k} = \lambda Y_k + (1-\lambda)E_{\text{Geo},k-1},$$

where λ is a smoothing or tuning parameter satisfying $0 < \lambda \leq 1$, and $E_{\text{Geo},0}$ is set to be

$$E_{\text{Geo},0} = 1/p_0.$$

For detecting increases in p, since values of Y_k's decrease as p increases, only a lower control limit is needed. The exact and approximate geometric lower control limits are given by

$$h_{\mathrm{GeoE}} = \frac{1}{p_0} - w_{\mathrm{Geo}} \frac{1}{p_0} \sqrt{\frac{(1-p_0)\lambda}{2-\lambda}(1-(1-\lambda)^{2k})},$$

and

$$h'_{\mathrm{GeoE}} = \frac{1}{p_0} - w_{\mathrm{Geo}} \frac{1}{p_0} \sqrt{\frac{(1-p_0)\lambda}{2-\lambda}},$$

respectively, where $k = 1, 2, \ldots$, and w_{Geo} is a multiplier that is determined so that a desired in-control performance can be achieved. The chart signals if $E_{\mathrm{Geo},k} < h_{\mathrm{GeoE}}$, provided that h_{GeoE} is used as the control limit, or $E_{\mathrm{Geo},k} < h'_{\mathrm{GeoE}}$ if the approximate limit is used instead.

If the primary interest is to detect increases in p, then only one control limit is needed – we use an upper control limit for monitoring Bernoulli observations and a lower control limit for monitoring geometric observations. In this situation, a reset is always recommended to be implemented in the EWMA chart. Particularly, for the Bernoulli case, the EWMA statistic with reset takes the form

$$E'_{\mathrm{Bern},k} = \max\left(p_0, \lambda X_k + (1-\lambda) E'_{\mathrm{Bern},k-1}\right).$$

Since in this dissertation performance comparisons are based on the steady state results, a reset is always used for all the EWMA charts being considered.

For more discussion on designing and applying the EWMA charts to monitor Bernoulli processes, see, for example, Sun and Zhang (2000), Spliid (2010), and Weiβ and Atzumüller (2010).

2.5 GLR Control Charts

In practical applications the shift size is always unknown in advance. Thus, due to lack of knowledge about the actual level of the shift in the process, many control charts perform poorly if the true shift size is substantially different from the size that the chart is designed to detect. For example, Shewhart charts work well exclusively for large shifts, while CUSUM charts are

known to have fast detection when the true shift size is close to p_1. It is rather challenging for a control chart to perform well for a wide range of shifts.

One approach to achieve this goal is to apply multiple charts simultaneously, each tuned to detect a particular shift size. For example, multiple CUSUM charts may be used together as a combination so that small, intermediate as well as large shifts could all be detected. The drawback is that its application usually requires specifications of too many control chart parameters. To overcome this, we could use an estimated size of shift that is evaluated presumably at every possible time of change and gives the maximum likelihood. Based on this estimated shift size we may construct a modified likelihood ratio test, which is developed for testing H_0, that the process parameter stays unchanged, against H_1, that the parameter has shifted to some unknown out-of-control value estimated by the maximum likelihood estimator. The control statistic formulated this way is essentially a GLR statistic.

The idea of applying generalized likelihood ratio tests in process change-point detection was proposed back in the 1970s. Lorden (1971) extended Page's (1954) CUSUM scheme to the form that involved the maximum likelihood estimates of both the change-point and the size of the change. This is usually referred to as the GLR scheme. The GLR chart has two major advantages over the other control charts in that, first of all, it has a maximum likelihood estimator as a "built-in" change-point estimator that could be used to provide important information about when a change has actually occurred in the process. Moreover, it is a generalization of the CUSUM scheme in that rather than only maximizing at a particular shift size p_1, maximization of the log-likelihood ratio is now taken over all possible values of the change-point and the process parameter, so it could be used to detect a wide range of shift sizes.

However, as is pointed out by Basseville and Nikiforov (1993), "practical implementation of the GLR algorithm is not always possible" at earlier times, because the amount of computations required to calculate the GLR statistic grows dramatically, since maximization of the log-likelihood ratio must be carried out for each possible value of the change-point. In order to reduce the computational burden of the GLR algorithm, various methods have been recommended. See Basseville and Nikiforov (1993) for a comprehensive review of the simplifications and modifications for change-point detection algorithms. Lai (1995)

also gave an excellent summary that included several ideas underlying the development of these procedures.

One option that is widely used now was introduced by Willsky and Jones (1976) in their paper regarding detection and estimation of jumps of unknown magnitudes in stochastic linear systems in the aerospace domain, where they proposed implementing the GLR scheme as a "finite memory filter" since maximizations of the log-likelihood ratio were only taken over a window of the past m samples. This algorithm largely reduced the computational complexity and is known as the "window-limited GLR scheme".

Willsky (1976) studied applications of the GLR procedure in the area of failure detection in dynamic systems. He mentioned that the GLR method can be used for detecting a wide range of failures, which would "provide an optimal decision rule for failure detection and useful failure identification information for use in system reorganization subsequent to the detection of a failure." He also suggested restricting the estimation of the change-point to a window of size m, which makes the algorithm more applicable, but with a reduction in accuracy of the parameter estimator.

Basseville and Benveniste (1983) presented another modified version of the GLR procedure, which involved mixing Willsky's GLR detector and Hinkley's CUSUM test, and showed that this algorithm was far less time consuming while still efficient and robust. In addition, by using normal approximations under local alternatives in conjunction with the instrumental variables method, Rougée et al. (1987) estimated the autoregressive parameters in a multivariate ARMA model to detect changes in the vibrating characteristics of mechanical systems. Furthermore, Benveniste et al. (1987) extended the idea from the instrumental variables method to the more general recursive algorithms for parameter estimation. Their methods were based on using non-likelihood procedures, which could be more convenient to execute than the GLR algorithms.

To detect a mean shift at some unknown instant in normal processes, Hawkins (1977) derived the likelihood ratio test statistics and their distributions in both cases when variance is known and unknown by using a recursion procedure, followed by some asymptotic results and an example concerning whether the stock exchange market had changed posture during an

interval. However, it was pointed out by Worsley (1979) that the null distribution for the situation with unknown variance in Hawkins' article was incorrect, and the correct version was provided in his paper.

Siegmund and Venkatraman (1995) investigated detecting a change in the normal mean with known variance as a special case in sequential change-point detection using the GLR statistic. They derived the asymptotic average run length under both the null and alternative hypotheses as the control limit approaches infinity, presented some numerical results using Monte Carlo simulations, and showed that the GLR chart is better than a standard CUSUM chart for detecting small and large changes, but there are only minor differences between the GLR method and the combined Shewhart-CUSUM charts. Some other situations such as detection for changes in a normal mean when the variance is unknown, or when the data are multivariate or autocorrelated were discussed in their paper as well.

Hawkins et al. (2003) studied detection and diagnosis of step changes from a normal process based on an unknown-parameter change-point formulation that does not require knowledge of the in-control parameters. Three scenarios were discussed: in the first scenario one assumes all the parameters except the change-point are known, and a CUSUM chart is appropriate in this situation; in the second scenario one assumes that the in-control mean and variance are known, but the magnitude of change in mean is unknown, where the GLR approach is suitable and could be applied to estimate the shift size; the last scenario is for the situation that none of the parameters are known, where they proposed using a two-sample t test between the left and right sections of the sequence, with maximizations taken across all possible change-points, and found the corresponding control limit using the Bonferroni inequality. They concluded that the performance of the GLR chart in change-point detection cannot quite match that of the CUSUM chart at some particular shift, but the GLR chart is more robust since it is close to optimal for a wide range of shifts.

Reynolds and Lou (2010) considered the problem of monitoring the mean of a normal process for a wide range of shift sizes. They evaluated the performance of the GLR control chart for sustained shifts, transient shifts, and drifts in the mean, with the likelihood ratio computed for sustained shift using a moving window of past observations. They found that the overall performance of the GLR chart is at least as good as the other options, such as combinations of

Shewhart and CUSUM charts and an adaptive CUSUM chart. The GLR chart they proposed did not require users to specify the values of any control chart parameters other than the moving window size and the control limit, for which specific recommendations were given in their paper. In addition, they provided an approximation to GLR chart using a finite set of CUSUM charts, which could be applied to simulate the performance of GLR chart for research purposes.

In statistical process control (SPC) literature, much of the research has focused on how quickly control charts signal the presence of a special cause variation in a process. On the other hand, as is pointed out by Pignatiello and Samuel (2001), "relatively little has been published on identification and diagnosis of the special cause variation after a signal is detected". This is due to the fact that this kind of diagnosis requires process-specific knowledge and each individual process is unique.

When a control chart signals, it is an indication to the process engineers that a process change has occurred, but it does not directly provide them with information about the time when this change actually happened. Nevertheless, knowing the process change-point is very helpful in assisting the process engineers in their search for the special cause variation and thus reduces the probability of incorrect process diagnosis. In this sense, process change-point estimation could lead to a quicker identification of the responsible special cause and a shorter process down time, which is important to all SPC applications.

Page (1954) proposed that for the CUSUM charts, we could use the starting point of the last SPRT, which is the point that begins to reject the null hypothesis of $p = p_0$, as the change-point estimator. The distribution of this estimator was studied by Hinkley (1970) and Nishina (1992). Srivastava (1994) further investigated this estimator and concluded that it was biased.

Similar to Page's idea, Nishina (1992) developed an estimator for the process change-point after an EWMA chart signals. He first defined a run on the EWMA chart as a sequence of points that are being plotted successively on one side of the center line. Then the change-point can be estimated by the starting point of the rejection run, which is the run that eventually exceeds the control limit and triggers a signal.

Another option for change-point estimation was introduced by Samuel et al. (1998a, 1998b), where they considered using the maximum likelihood estimator as an estimator for the

change-point and evaluated its performance when used after a signal given by the Shewhart $\bar{X}$-chart. Consequently, this estimator can be applied when an out-of-control signal is detected by any control chart monitoring a normal process mean, including the Shewhart $\bar{X}$, CUSUM and EWMA charts.

In fact, change-point estimation is automatically accomplished by the GLR chart, for the GLR statistic is maximized at each time point, and based on this we can easily identify the location where a change is most likely to have occurred. In addition, an estimated shift size is also provided, making the GLR chart a very competitive tool in the diagnosis after a signal is given.

With the development of modern computing technology, the computational burden involved in the GLR algorithm would no longer be a major obstacle for its applications in practice. To further reduce the burden of computing, we will build our GLR chart based on the same algorithm that was proposed by Willsky and Jones (1976) as the “window-limited GLR scheme”.

Chapter 3

Derivation and Plots of the Binomial GLR Chart

3.1 Derivation of the Binomial GLR Statistic

After sample k has been collected we have observed k independent binomial random variables $T_1, T_2, \dots, T_k$. We want to test the null hypothesis that there has been no change in p against the alternative hypothesis that the process proportion p has increased from p_0 to some unknown $p_1$between samples τ and $\tau + 1$, for some $\tau < k$. The likelihood function at sample k under the alternative hypothesis is

$$L(\tau, p_1|T_1, T_2, \dots, T_k) = \left\{\prod_{i=1}^{\tau} \binom{n}{T_i} p_0^{T_i}(1-p_0)^{n-T_i}\right\}\left\{\prod_{i=\tau+1}^{k} \binom{n}{T_i} p_1^{T_i}(1-p_1)^{n-T_i}\right\}.$$

At a given value of τ we estimate p_1 by the maximum likelihood estimator

$$\hat{p}_1 = \max\left\{p_0, \frac{\sum_{i=\tau+1}^{k} T_i}{(k-\tau)n}\right\}, \tag{3.1}$$

where the restriction $\hat{p}_1 \geq p_0$ is only needed for the one-sided case of detecting increases in p. Without such a restriction, we would have a two-sided chart that signals either $\hat{p}_1$ is too small or $\hat{p}_1$ is too large. The likelihood function at sample k under the null hypothesis that p remains unchanged is

$$L(\infty, p_0|T_1, T_2, \dots, T_k) = \prod_{i=1}^{k} \binom{n}{T_i} p_0^{T_i}(1-p_0)^{n-T_i}.$$

To determine if there is indeed an increase in p, a binomial GLR control statistic $\lambda_{\text{Bin},k}$ is constructed based on the following log-likelihood ratio:

$$\lambda_{\text{Bin},k} = ln\frac{\max\limits_{0\leq\tau<k,\ p_0\leq p_1<1} L(\tau, p_1|T_1, T_2, \dots, T_k)}{L(\infty, p_0|T_1, T_2, \dots, T_k)}$$

$$= \ln \frac{\max\limits_{0\le\tau<k,\, p_0\le p_1<1} \left\{\prod_{i=1}^{\tau}\binom{n}{T_i} p_0^{T_i}(1-p_0)^{n-T_i}\right\}\left\{\prod_{i=\tau+1}^{k}\binom{n}{T_i} p_1^{T_i}(1-p_1)^{n-T_i}\right\}}{\prod_{i=1}^{k}\binom{n}{T_i} p_0^{T_i}(1-p_0)^{n-T_i}}$$

$$= \ln \frac{\max\limits_{0\le\tau<k,\, p_0\le p_1<1} \prod_{i=\tau+1}^{k}\binom{n}{T_i} p_1^{T_i}(1-p_1)^{n-T_i}}{\prod_{i=\tau+1}^{k}\binom{n}{T_i} p_0^{T_i}(1-p_0)^{n-T_i}}$$

$$= \max_{0\le\tau<k} n(k-\tau)\left[\hat{p}_1 \ln\frac{\hat{p}_1(1-p_0)}{p_0(1-\hat{p}_1)} + \ln\frac{1-\hat{p}_1}{1-p_0}\right]. \tag{3.2}$$

An estimate of the change-point $\hat{\tau}$ is then given by

$$\hat{\tau} = \operatorname*{argmax}_{0\le\tau<k} n(k-\tau)\left[\hat{p}_1 \ln\frac{\hat{p}_1(1-p_0)}{p_0(1-\hat{p}_1)} + \ln\frac{1-\hat{p}_1}{1-p_0}\right].$$

Notice that if the samples after τ consist of items that are all nonconforming, then $\sum_{i=\tau+1}^{k} T_i$ would be equal to $(k-\tau)n$, so there is a chance that $\hat{p}_1 = 1$. In particular, consider a shift occurring between the last two samples (sample $k-1$ and sample k). If n is very small, say, less than 5, and the out-of-control proportion is large, it may happen that $\sum_{i=\tau+1}^{k} T_i = n$, in which case $\hat{p}_1 = 1$, which will lead to an undefined GLR statistic. But as long as n is not too small and the shift is not too close to 1, the probability of having $\hat{p}_1 = 1$ is extremely small. In this dissertation, the sample sizes being considered for the binomial charts are at least 10, so we can neglect the possibility that $\hat{p}_1 = 1$.

If all past samples are examined and we maximize $\lambda_{\text{Bin},k}$ over $0 \le \tau < k$ at each sampling point, it could be time consuming to obtain the GLR statistic when k is very large. Instead, maximization can be performed only for a window of the past m samples (see Willsky and Jones (1976)). In this case, the binomial GLR statistic is modified as

$$\lambda_{\text{Bin},m,k} = \begin{cases} \max\limits_{0\le\tau<k} n(k-\tau)\left[\hat{p}_1 \ln\frac{\hat{p}_1(1-p_0)}{p_0(1-\hat{p}_1)} + \ln\frac{1-\hat{p}_1}{1-p_0}\right], & k = 1, 2, \ldots, m \\ \max\limits_{k-m\le\tau<k} n(k-\tau)\left[\hat{p}_1 \ln\frac{\hat{p}_1(1-p_0)}{p_0(1-\hat{p}_1)} + \ln\frac{1-\hat{p}_1}{1-p_0}\right], & k = m+1, m+2, \ldots \end{cases}. \tag{3.3}$$

The window size m will be chosen to be sufficiently large that the performance of the GLR chart with the window is essentially the same as the performance of the GLR chart without the window (it will be shown later that choosing m in this way gives the best overall performance).

The binomial GLR chart signals when $\lambda_{\text{Bin},m,k} > h_{\text{GLR}}$, where h_{GLR} is a pre-determined control limit chosen to achieve the desired in-control performance.

3.2 Example Plots of the Binomial GLR Chart

As an example, we first simulated 100 samples from the binomial distribution with $p_0 = 0.01$ and $n = 100$ to show what the data and the binomial GLR chart look like when the process is in-control. After the first 100 samples, 40 more samples were generated from the same process but with p increased so that the process is now regarded as out-of-control (with $\tau = 100$). In particular, two out-of-control cases were considered. One case was for a small shift with $p = 0.015$ and $n = 100$; another case considers a moderate shift with $p = 0.03$ and $n = 100$.

For the first case, the simulated values of T_k are plotted in Figure 3.1, the binomial GLR chart is displayed in Figure 3.2, and Figures 3.3 and 3.4 are two binomial CUSUM charts with different tuning parameters – one tuned at 0.015 and the other one tuned at 0.03. Figure 3.5 represents a trace plot of the process change-point estimator $\hat{\tau}$ from the binomial GLR chart.

In Figure 3.1, the control limit of the Shewhart chart was chosen to be $h'_{\text{S}} = 5$, and this gives an in-control ANSS value of 291.35. Note that the Shewhart chart signals only when $T_k \geq h'_{\text{S}}$, so there are no signals (false-alarms) in the first 100 in-control observations. In Figure 3.2, the control limit of the binomial GLR chart is set at $h_{\text{GLR}} = 4.13$ to give an in-control ANSS value of 293.50, which is close to the in-control ANSS value of the Shewhart chart. The limits for the two CUSUM charts in Figure 3.3 and 3.4 are found in a similar manner.

Since the shift size considered here is very small, it is not surprising that the Shewhart chart in Figure 3.1 is not able to detect this shift within 40 out-of-control samples. Based on recommendations given later on the design of the binomial GLR chart, a moving window of size $m = 300$ is used; nevertheless, since we only simulated 140 samples in total here, there was actually no window used for the GLR chart in Figure 3.2. The shift shows up more clearly on the binomial GLR chart than on the Shewhart chart, with a signal given at sample 114, for which the corresponding GLR statistic is 4.147 and the estimated shift size is $\hat{p}_1 = 0.0177$. In this example,

the binomial GLR chart signals much faster than the Shewhart chart, which again illustrates that the latter is ineffective in detecting small shifts.

The CUSUM chart in Figure 3.3 with $p_1 = 0.015$ (which equals to the true shift size in the example) signaled at sample 114; however, the other CUSUM chart in Figure 3.4 with $p_1 = 0.03$ was not able to detect this shift due to the relatively large value of the tuning parameter. The plot of the GLR statistic looks very similar to that of the CUSUM statistic with $p_1 = 0.015$, because the value of $\hat{p}_1$ estimated in the GLR chart is close to the value of p_1 specified in the CUSUM chart. Both of them out-performed the CUSUM chart with the tuning parameter $p_1 = 0.03$.

Values of $\hat{\tau}$ were calculated after each sample and are plotted in Figure 3.5. For the first 100 observations, τ is estimated based on just the random in-control variation of the observations. After the shift occurred between sample 100 and 101, it is clearly shown on the plot that the maximum likelihood estimator picked up this change right away. Since the GLR chart detected this shift at sample 114, of particular interest is the value of $\hat{\tau}$ at that time, and which is equal to 100. This suggests that the process change is most likely to have occurred at around sample 100, which is actually the true process change-point.

FIGURE 3.1. Plot of 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.015$. No signal is given.

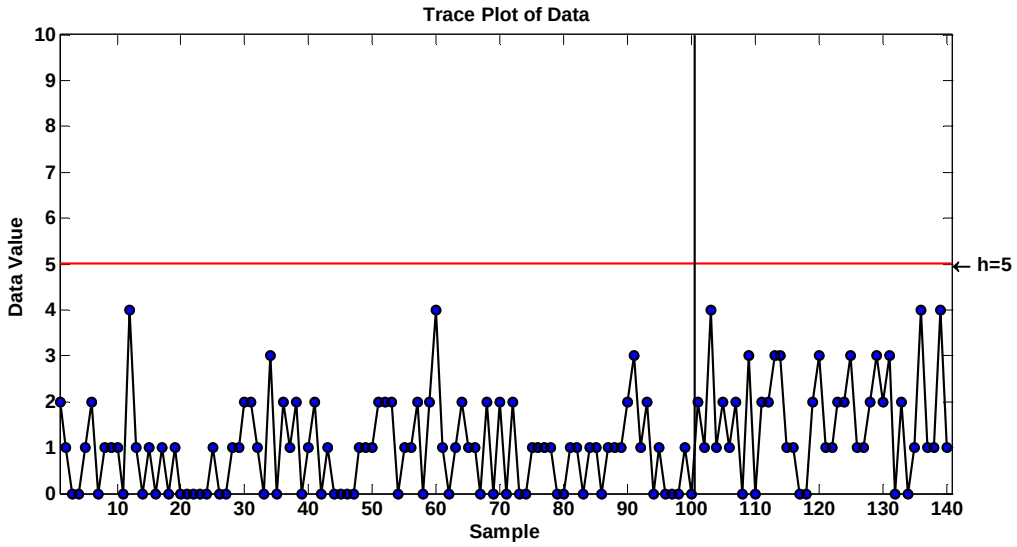


FIGURE 3.2. Plot of binomial GLR statistics for 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.015$. An out-of-control signal is given at sample 114 based on a control limit of 4.13.

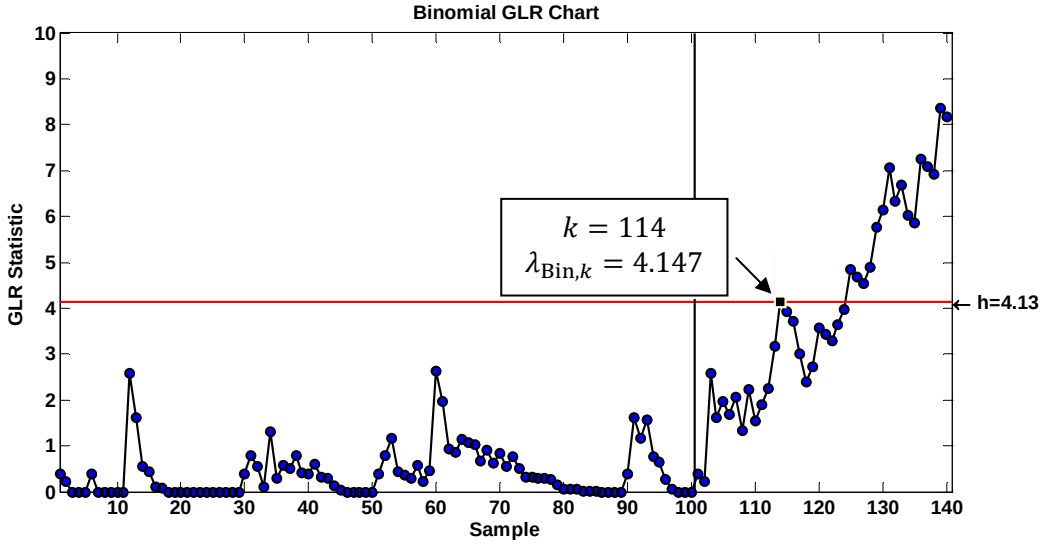


FIGURE 3.3. Plot of binomial CUSUM statistics tuned at $p_1 = 0.015$ for 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.015$. An out-of-control signal is given at sample 114 based on a control limit of 2.96.

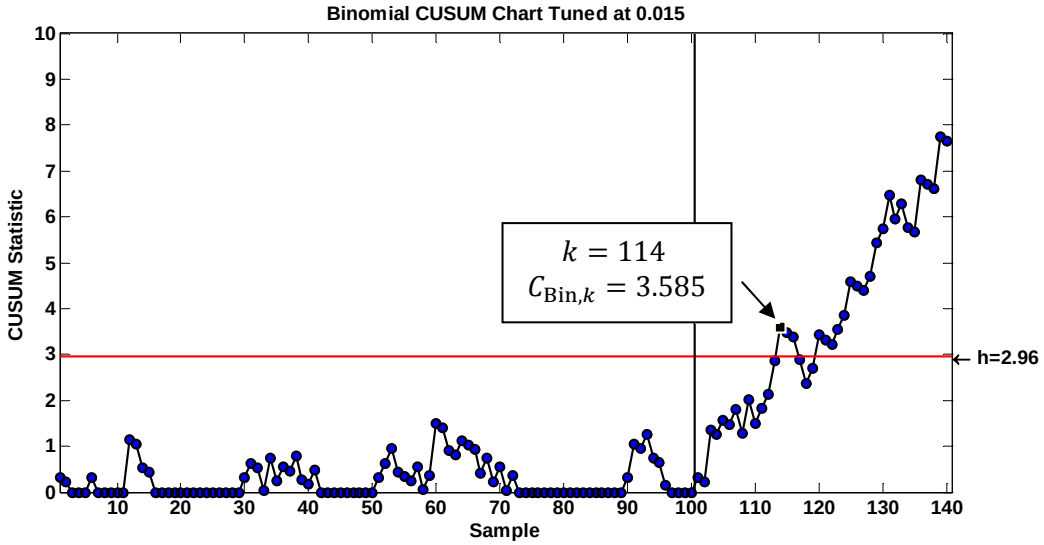


FIGURE 3.4. Plot of binomial CUSUM statistics tuned at $p_1 = 0.03$ for 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.015$. No signal is given.

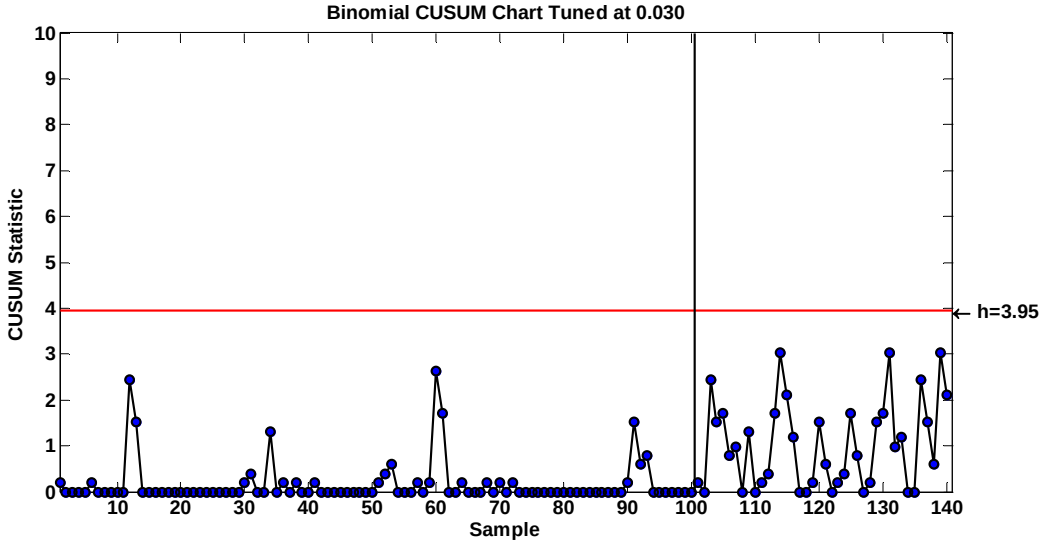
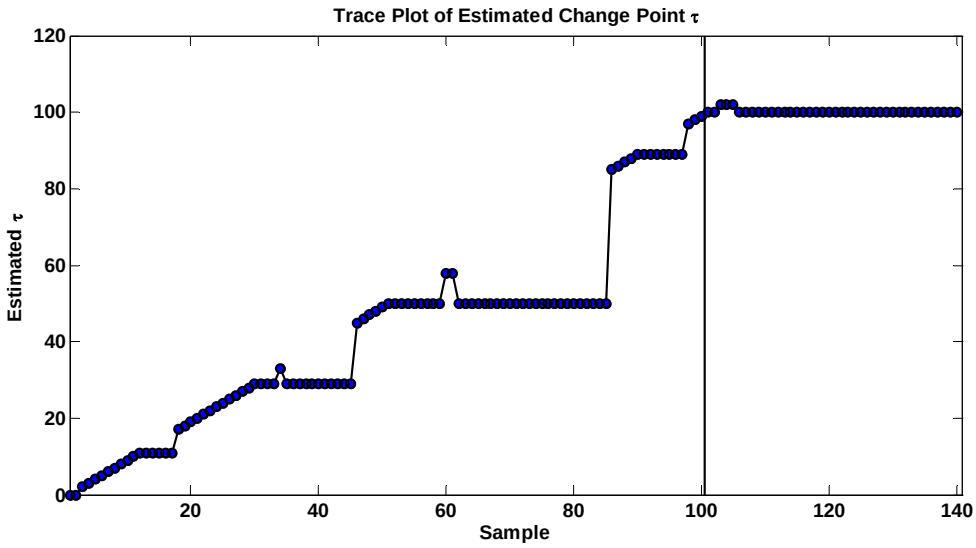


FIGURE 3.5. Trace plot of the estimated process change-point from the binomial GLR chart when $p = 0.015$.



Now consider a larger shift size of 0.03. The simulated values of T_k are plotted as a Shewhart np-chart in Figure 3.6, the binomial GLR chart is displayed in Figure 3.7, and Figures 3.8 and 3.9 are two binomial CUSUM charts with different tuning parameters – one tuned at 0.015 and the other one tuned at 0.03. Here, the control limit for each chart is the same as in the

first case, so that they all had the same in-control performance. Figure 3.10 represents a trace plot of the process change-point estimator $\hat{\tau}$ from the binomial GLR chart.

We see from Figure 3.6 that the Shewhart np-chart with $h'_S = 5$ detected the shift size of 0.03 within three samples. In Figure 3.7, the binomial GLR chart had a signal at sample 104. Figure 3.8 shows that the binomial CUSUM chart with $p_1 = 0.015$ was able to detect this shift fairly quickly, but one more sample was needed for it to signal when compared to the GLR chart. Figure 3.9 shows that the CUSUM chart with $p_1 = 0.03$ seemed to have a similar performance as the binomial GLR chart, with the signal time being the same as that of the GLR chart. The estimated change-point given by Figure 3.10 indicates that a shift occurred around sample 101 in this process.

FIGURE 3.6. Plot of 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.03$. An out-of-control signal is given at sample 103 based on a control limit of 5.

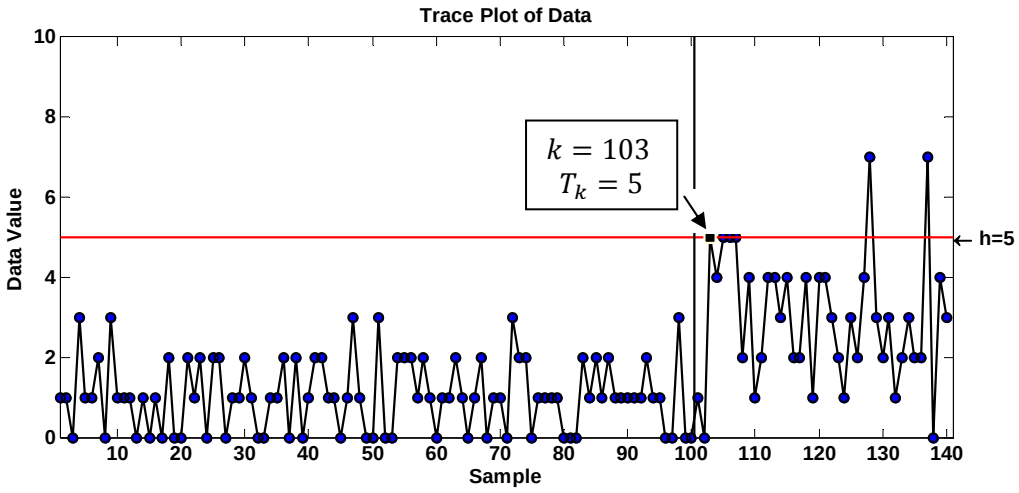


FIGURE 3.7. Plot of binomial GLR statistics for 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.03$. An out-of-control signal is given at sample 104 based on a control limit of 4.13.

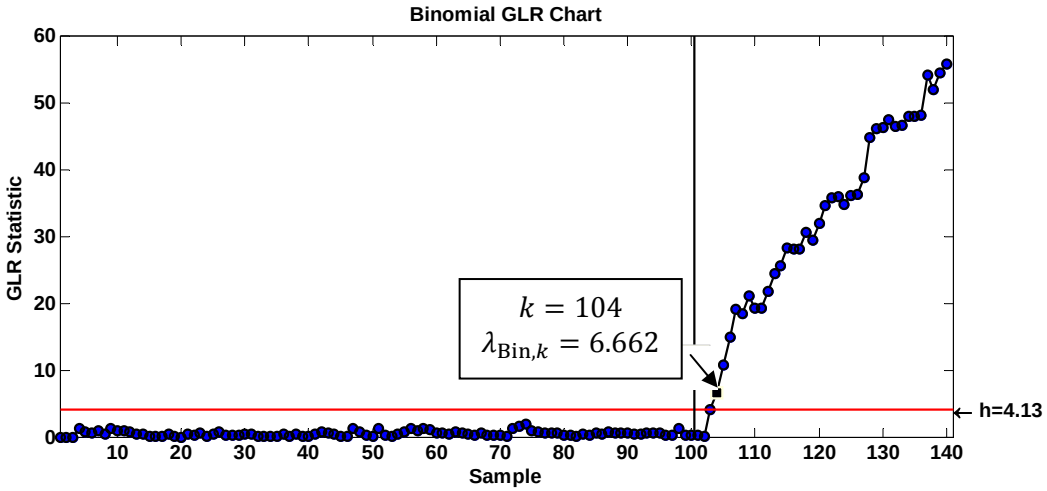


FIGURE 3.8. Plot of binomial CUSUM statistics tuned at $p_1 = 0.015$ for 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.03$. An out-of-control signal is given at sample 105 based on a control limit of 2.96.

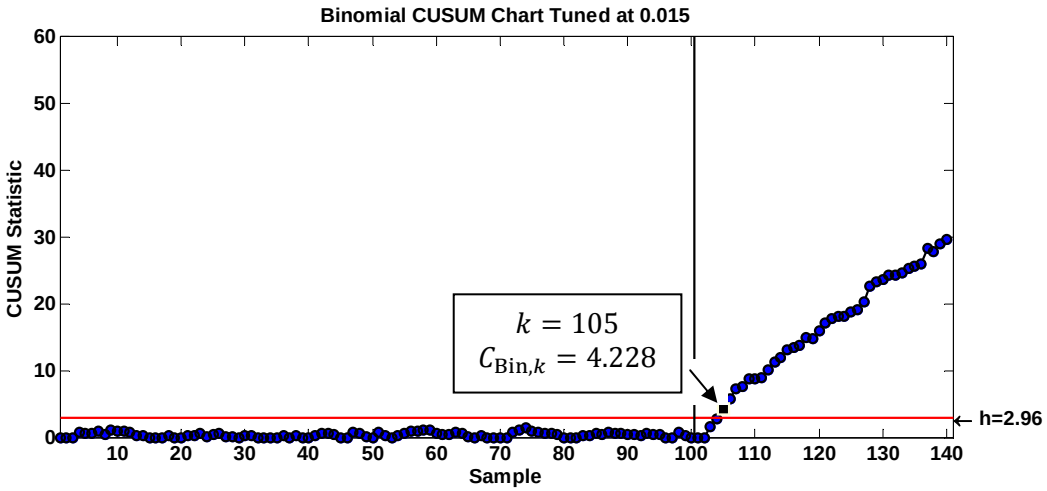


FIGURE 3.9. Plot of binomial CUSUM statistics tuned at $p_1 = 0.03$ for 100 in-control values of T_k with $p_0 = 0.01$ and $n = 100$, and 40 out-of-control values of T_k with $p = 0.03$. An out-of-control signal is given at sample 104 based on a control limit of 3.95.

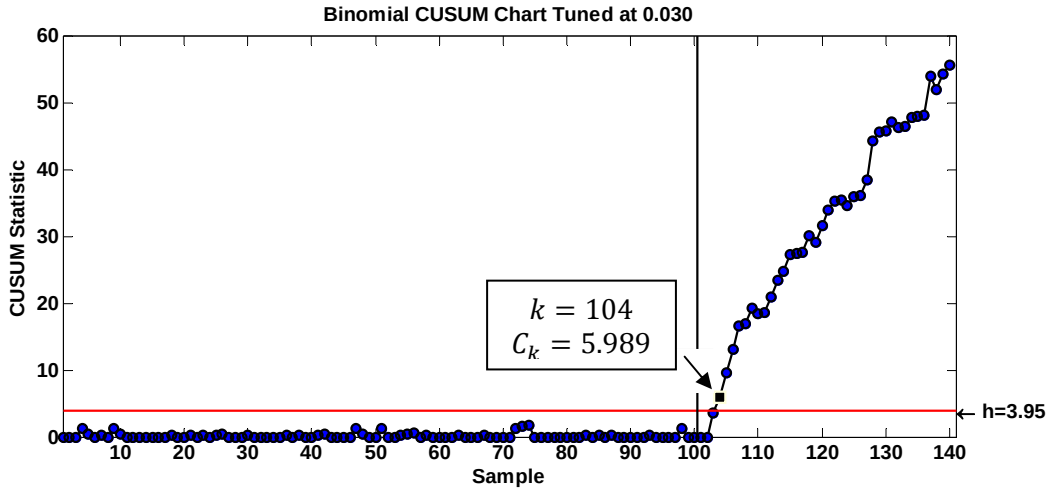
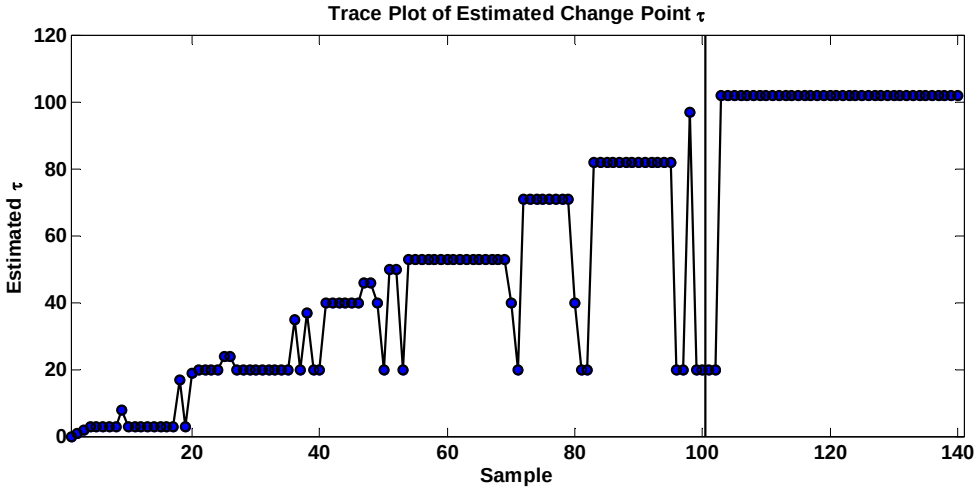


FIGURE 3.10. Trace plot of the estimated process change-point from the binomial GLR chart when $p = 0.03$.



Chapter 4

Designing the Binomial GLR Chart

Unlike the binomial CUSUM chart, the binomial GLR chart does not require specification of the out-of-control value p_1. Instead, only the moving window size m and the control limit h_{GLR} are needed for designing a binomial GLR chart.

4.1 Selecting the Window Size

The performance of a binomial GLR chart will be affected by the value of p_0 and the choice of the sample size n. Therefore, when we evaluate the effect of the window size m on the performance of the GLR chart, we considered different values of p_0 and different sample sizes with the same p_0. The goal is to find values of m so that the resulting performance is very close to the performance given by not using the window.

For the case of $p_0 = 0.01$ and $n = 100$, Table 4.1 gives in-control ANSS values and out-of-control SSANSS values for various shifts in p with m ranging from 10 to 800. The control limits for all the GLR charts with different window sizes are set to be the same here at $h_{\text{GLR}} = 4.13$, which leads to slightly different in-control ANSS values for each GLR chart. Due to the discreteness of the binomial distribution, using $h_{\text{GLR}} = 4.13$ gives in-control ANSS values for the values of m that are as close as we can get to the in-control ANSS value of 291.35 for the Shewhart chart with $h'_{\text{S}} = 5$. The SSANSS values are obtained assuming that $\tau = 100$.

It is apparent from Table 4.1 that a large value of m is required for good overall performance because a large value of m is always needed to quickly detect small shifts in p. It appears that a window of size $m = 300$, which is approximately the same as the desired in-control ANSS value (291.35), is sufficiently large for detecting shifts as small as 0.013. This is equivalent to keeping the last 300 samples or 30,000 individual observations. Although we have data from 30,000 observations in the moving window, we actually only need to store 300 values of T_k, so storage requirements are still reasonable. In fact, it appears that a smaller value of m

may be adequate as well, but using $m = 300$ is definitely large enough even for a very small shift. In addition, there seems to be essentially no difference for various values of m on the out-of-control performance if the shift size is around 0.05 or larger.

TABLE 4.1. Effect of the window size on in-control ANSS and out-of-control SSANSS values for the binomial GLR chart when $p_0 = 0.01$, $n = 100$, $\tau = 100$, and $h_{\mathrm{GLR}} = 4.13$.

p= \ *m*=	10	50	100	200	300	400	800
0.010	331.85	296.46	293.71	293.76	294.04	293.01	293.29
0.013	64.68	49.77	47.80	47.41	47.37	47.12	47.19
0.014	43.58	33.68	32.57	32.57	32.60	32.54	32.60
0.015	31.02	24.53	24.11	24.08	24.08	24.11	24.11
0.017	18.15	15.35	15.23	15.29	15.30	15.24	15.24
0.020	10.34	9.60	9.53	9.53	9.52	9.51	9.53
0.025	5.98	5.85	5.83	5.83	5.84	5.82	5.84
0.030	4.33	4.29	4.29	4.29	4.27	4.30	4.28
0.035	3.47	3.46	3.45	3.46	3.46	3.47	3.46
0.040	2.97	2.96	2.95	2.96	2.96	2.96	2.96
0.050	2.38	2.38	2.38	2.38	2.38	2.38	2.39

TABLE 4.2. Effect of the window size on in-control ANSS and out-of-control SSANSS values for the binomial GLR chart when $p_0 = 0.01$, $n = 16$, $\tau = 1{,}000$, and $\mathrm{h}_{\mathrm{GLR}} = 6.29$.

p= \ *m*=	100	500	1000	1500	2000	2500	5000
0.010	1996.01	1965.82	1963.65	1963.61	1959.49	1958.89	1958.56
0.013	489.16	413.20	399.63	395.12	395.02	393.90	394.43
0.014	331.87	273.27	265.55	265.79	265.90	265.84	265.94
0.015	232.98	193.62	192.95	192.75	192.27	192.17	191.75
0.017	130.52	115.75	115.93	115.65	116.36	115.78	116.00
0.020	70.11	67.48	67.12	67.32	67.47	67.11	67.20
0.025	36.68	36.44	36.50	36.49	36.48	36.46	36.52
0.030	24.10	23.94	23.89	23.98	23.93	23.94	23.82
0.035	17.55	17.53	17.47	17.45	17.51	17.49	17.51
0.040	13.72	13.71	13.75	13.72	13.72	13.71	13.72
0.050	9.50	9.49	9.48	9.49	9.49	9.49	9.48

Results are presented in Table 4.2 for the binomial GLR chart with the same in-control proportion $p_0 = 0.01$ as in Table 4.1, but now we consider a different sample size, $n = 16$. In this case a Shewhart chart with $h'_{\mathrm{S}} = 2$ gives an in-control ANSS of only 91.47. This value seems to be unreasonably small, so the limit $h'_{\mathrm{S}} = 3$ is used, which gives an in-control ANSS value of 1,968.73. The GLR control limit is then chosen to be $h_{\mathrm{GLR}} = 6.29$ to ensure that the in-

control ANSS values for the values of m are close to 1,968.73. SSANSS values are obtained for shifts in p with m ranging from 10 to 5,000, assuming that $\tau = 1{,}000$. Table 4.2 shows that in order to detect shifts as small as 0.013, it is sufficient to use a moving window of size $m = 2{,}000$, which requires storing the most recent 2,000 samples. Notice that this value of m is again very close to the target in-control ANSS value (1,968.73), which implies that to design a binomial GLR chart, we could simply set the window size m to be close to the desired in-control ANSS value.

The same conclusions about the choice of m can be drawn from processes with different values of p_0 and n (results not shown here). In fact the general relationship between m and the in-control ANSS and shift size that we found through simulation has been confirmed theoretically by Lai (1998), where he extended Lorden's (1971) finding on the asymptotic properties of the GLR chart without the window and showed that the GLR chart with the window is also asymptotically optimal if the window size $m = O(\ln\gamma / I_{\min})$, where γ is the in-control ANSS and $I_{\min}$ is the Kullback-Leibler information number corresponding to the minimum shift size of interest. From here, we see both theoretically and numerically that a large m is needed if the desired in-control ANSS is large or the target shift size is small, and a small m is sufficient when the in-control ANSS is small or the expected shift size is large.

In Table 4.3, we consider the effect of using different window sizes for combinations of three sample sizes and two in-control ANOS values, all from processes with $p_0 = 0.001$. Both the in-control case and an out-of-control shift of 0.0015 are simulated, with the length of the in-control period used to reach steady state for each combination chosen to be one half of the corresponding in-control ANSS value. The results confirmed our recommendation that we should select the window size to match the in-control ANSS value. As an example, consider the sample size $n = 1{,}000$ and in-control ANOS approximately equal to 100,000. If we use a moving window of size $m = 100$, the in-control ANOS value is 101.00 and the out-of-control ANOS value at shift $p = 0.0015$ is 15.69. For a moving window 10 times larger, i.e. $m = 1{,}000$, the corresponding ANOS values are 101.08 and 15.78, respectively, which are extremely close to the results given by using a smaller window size $m = 100$.

Table 4.4 and 4.5 gives similar results for processes of $p_0 = 0.01$ and $p_0 = 0.1$, respectively. In Table 4.4, the combination of $n = 10$ and in-control ANOS $= 100{,}000$ is omitted, for an in-control ANOS value this large is unlikely to occur in practice. No matter which combination of the sample size and in-control ANOS value is considered, the same conclusion holds for selection of the window size: we should pick an m value that is close to the in-control ANSS value.

Based on the simulation results for different values of p_0 and n, we recommend that the window size m be chosen close to the desired in-control ANSS value so that it is sufficiently large to give performance that is virtually identical to the case in which no window is used. Thus, to select a reasonable window size m, all we need to know is the desired in-control ANSS value, or the desired in-control ANOS value and the corresponding sample size n, from which the ANSS value can be easily calculated.

TABLE 4.3. Effect of different window sizes on SSANSS values of the binomial GLR chart when $p_0 = 0.001$, based on various combinations of sample sizes and in-control ANOS values.

n=100, In-Control ANOS=10000, h_{GLR}=2.82							
m	10	20	50	100	200	500	1000
p=0.0010	122.59	121.76	120.74	120.16	119.83	120.00	120.27
p=0.0015	52.03	51.10	50.26	50.06	50.06	49.76	50.30
n=100, In-Control ANOS=100000, h_{GLR}=5.35							
m	100	200	500	1000	1500	2000	5000
p=0.0010	1202.76	1197.77	1190.52	1188.99	1191.08	1188.54	1188.11
p=0.0015	232.33	219.22	209.30	208.04	207.88	207.35	207.47
n=500, In-Control ANOS=10000, h_{GLR}=1.29							
m	8	10	15	20	30	50	100
p=0.0010	17.55	17.57	17.57	17.51	17.56	17.56	17.57
p=0.0015	8.15	8.06	8.06	7.98	8.03	8.04	8.06
n=500, In-Control ANOS=100000, h_{GLR}=3.89							
m	50	100	150	200	300	500	1000
p=0.0010	198.23	197.67	197.66	197.34	197.19	197.33	197.05
p=0.0015	34.74	34.20	34.16	34.18	34.08	34.15	34.19
n=1000, In-Control ANOS=10000, h_{GLR}=0.77							
m	4	6	8	10	20	50	100
p=0.0010	9.19	9.18	9.20	9.20	9.18	9.20	9.19
p=0.0015	4.40	4.39	4.39	4.38	4.40	4.36	4.39
n=1000, In-Control ANOS=100000, h_{GLR}=2.86							
m	10	25	50	100	200	500	1000
p=0.0010	106.95	101.71	101.18	101.00	100.94	100.85	101.08
p=0.0015	17.57	17.57	15.83	15.69	15.79	15.79	15.78

TABLE 4.4. Effect of different window sizes on SSANSS values of the binomial GLR chart when $p_0 = 0.01$, based on various combinations of sample sizes and in-control ANOS values.

n=10, In-Control ANOS=10000, h_{GLR}=5.53							
m	100	200	500	1000	1500	2000	5000
p=0.010	835.65	831.42	829.76	829.53	829.73	828.63	829.89
p=0.013	337.66	328.04	319.29	317.28	316.77	316.39	316.08
n=50, In-Control ANOS=10000, h_{GLR}=3.89							
m	50	100	150	200	300	500	1000
p=0.010	205.19	204.40	204.23	203.07	203.39	203.90	202.66
p=0.013	61.58	60.34	59.67	59.58	59.67	59.95	59.68
n=50, In-Control ANOS=100000, h_{GLR}=6.49							
m	500	1000	1500	2000	3000	5000	10000
p=0.010	1996.91	1996.63	1993.45	1990.81	1991.62	1983.92	1983.89
p=0.013	178.56	177.14	177.09	177.23	177.23	177.14	177.39
n=100, In-Control ANOS=10000, h_{GLR}=2.89							
m	10	20	50	100	200	500	1000
p=0.010	110.78	105.77	104.34	104.10	104.21	104.15	104.05
p=0.013	31.18	28.63	27.53	27.54	27.42	27.31	27.37
n=100, In-Control ANOS=100000, h_{GLR}=5.87							
m	100	200	500	1000	1500	2000	5000
p=0.010	909.42	903.76	901.54	900.36	901.24	899.86	898.82
p=0.013	87.92	82.95	81.89	81.77	81.76	81.80	81.89

TABLE 4.5. Effect of different window sizes on SSANSS values of the binomial GLR chart when $p_0 = 0.1$, based on various combinations of sample sizes and in-control ANOS values.

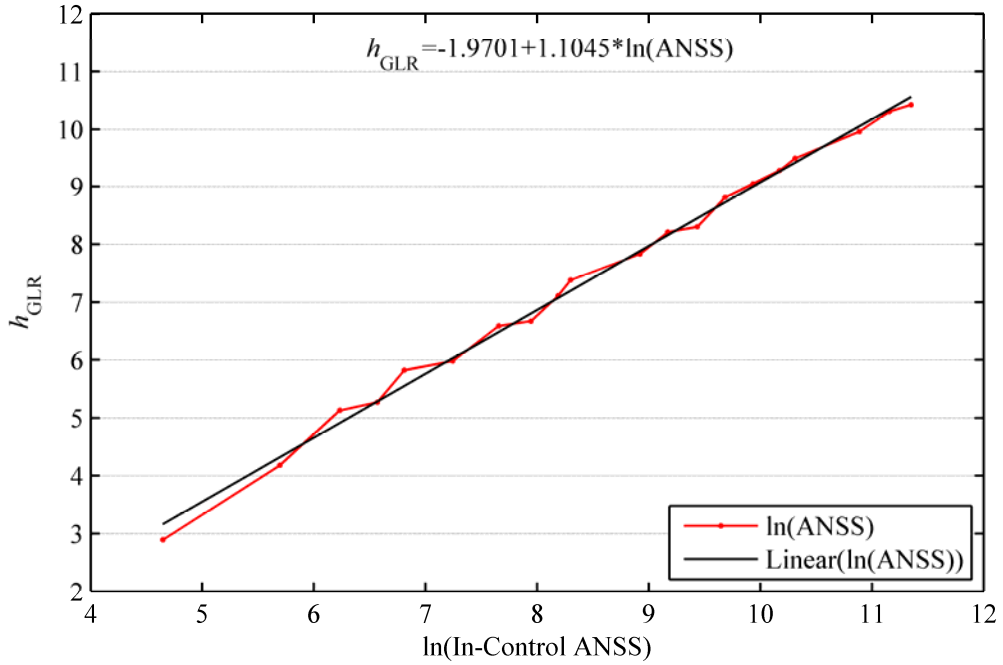
n=10, In-Control ANOS=1000, h_{GLR}=3.12							
m	10	20	50	100	200	500	1000
p=0.10	125.21	118.61	118.48	117.75	116.66	116.69	117.21
p=0.11	73.94	68.21	66.28	65.96	66.05	65.85	66.03
n=10, In-Control ANOS=10000, h_{GLR}=6.03							
m	100	200	500	1000	1500	2000	5000
p=0.10	1996.91	1996.63	1993.45	1990.81	1991.62	1983.92	1983.89
p=0.11	392.98	371.15	350.55	343.92	342.27	342.50	342.55
n=50, In-Control ANOS=1000, h_{GLR}=1.48							
m	8	10	15	20	30	50	100
p=0.10	21.26	21.20	21.23	21.17	21.14	21.03	21.09
p=0.11	11.68	11.54	11.51	11.49	11.42	11.41	11.42
n=50, In-Control ANOS=10000, h_{GLR}=4.09							
m	50	100	150	200	300	500	1000
p=0.10	229.75	224.90	224.15	223.49	224.05	223.72	222.68
p=0.11	59.15	57.58	56.97	56.84	56.76	56.61	56.68
n=100, In-Control ANOS=1000, h_{GLR}=0.81							
m	4	6	8	10	20	50	100
p=0.10	9.81	9.79	9.78	9.77	9.76	9.74	9.74
p=0.11	5.54	5.50	5.48	5.46	5.47	5.49	5.48
n=100, In-Control ANOS=10000, h_{GLR}=3.44							
m	10	20	50	100	200	500	1000
p=0.10	102.92	99.43	99.48	98.58	98.65	96.89	97.62
p=0.11	29.82	27.48	26.03	25.76	25.88	25.80	25.85

4.2 Selecting the Control Limit

The selection of h_{GLR} depends on p_0, n, and the desired in-control ANSS value. In general, for a particular combination of p_0 and n, h_{GLR} is approximately linearly related to the natural log of the in-control ANSS values. To demonstrate this for the case of $p_0 = 0.01$, $n = 100$, and $m = 1{,}000$, we found the values of h_{GLR} for 20 in-control ANSS values ranging from 100 to 85,000, and plotted h_{GLR} versus the log of the ANSS in Figure 4.1. A fitted line of

h_{GLR} as a function of log(ANSS) is given as well. Figure 4.1 shows that although h_{GLR} and ln(ANSS) are not perfectly fitted by a straight line (this might be mainly due to the discreteness of the binomial distribution), there exists a strong linear relationship between them and the corresponding R^2 is above 0.99.

FIGURE 4.1. Fitted line plot of h_{GLR} and ln(in-control ANSS) for $p_0 = 0.01, n = 100$, and $m = 1{,}000$.



Intuitively, p_0 and n should each have an impact on the appropriate choice of h_{GLR}, but simulation results show that their influence on h_{GLR} is rather minor and even negligible. We selected 27 values of h_{GLR} that range from 3.5 to 10 with an increment of 0.25, and seven different combinations of p_0 and n, with p_0 ranging from 0.001 to 0.1 and n ranging from 10 to 1,000. Then we simulated the corresponding in-control ANSS values to illustrate the effects of p_0 and n on h_{GLR}. As is shown in Figure 4.2, the relationships between ln(ANSS) and h_{GLR} for different p_0 and n combinations have very similar patterns. If a straight line were fitted for each combination, the estimated parameters should only be slightly different from the results when the combinations are considered together. Furthermore, these differences are mainly due to the discreteness of the various binomial distributions, rather than due to any difference in the general

trend of the relationship between h_{GLR} and ln(ANSS). In this case, a relatively accurate h_{GLR} value could be easily obtained no matter what values p_0 and n are specified.

FIGURE 4.2. Plots of h_{GLR} and ln(in-control ANSS) for different p_0 and n combinations.

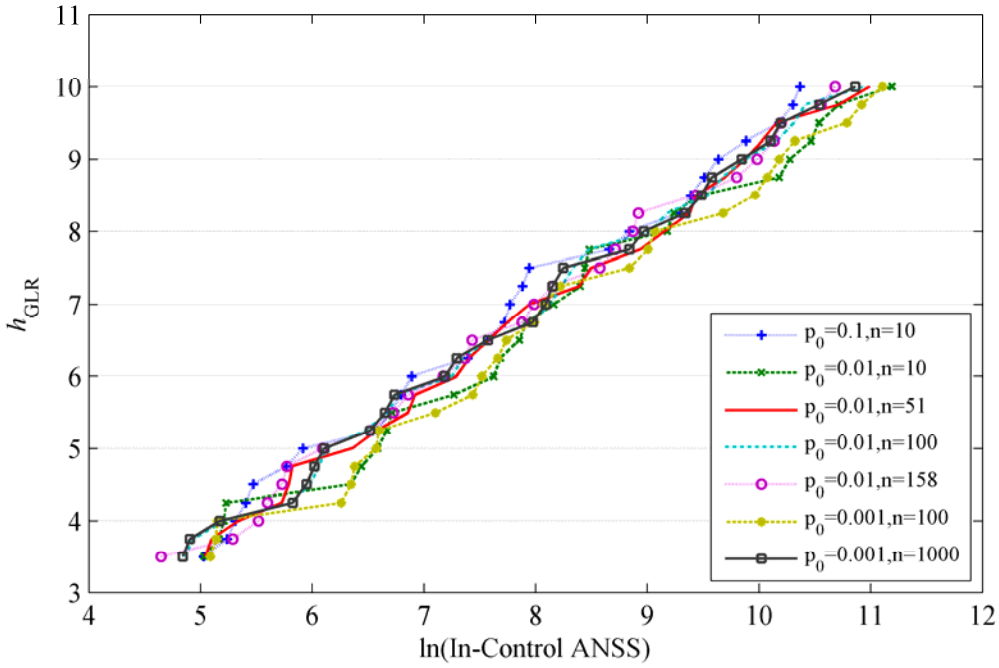


Table 4.6 gives numerical summaries from linear regression between h_{GLR} and ln(in-control ANSS) for each of the seven combinations from Figure 4.2, with the in-control ANSS values ranging from about 50 to 10,000. Higher order terms (quadratic, cubic, etc.) were also considered, but it didn't seem to be necessary to include any of them in the model. Also shown in this table is that from case to case, the estimated linear regression slopes are fairly close, but the intercepts are slightly different. An 'Average' model that uses the average value of the estimated parameters from seven individual models (shown at the bottom of the table) is

$$h_{\text{GLR}} = -2.133 + 1.128 \times \ln(\text{in-control ANSS}). \qquad (4.1)$$

A linear regression model between h_{GLR} and ln(in-control ANSS) based on the whole data set could also be used, and it gives results very similar to Equation (4.1).

TABLE 4.6. Linear regression results for different p_0 and n combinations.

			Linear Trend		R^2
p_0	n	m	Intercept	ln(In-Control ANSS)	Linear
0.001	100	1000	-2.500	1.147	0.986
0.001	1000	100	-2.053	1.123	0.991
0.010	10	1000	-2.493	1.154	0.983
0.010	51	200	-2.017	1.108	0.988
0.010	100	100	-2.028	1.122	0.991
0.010	158	80	-2.010	1.127	0.995
0.100	10	800	-1.828	1.115	0.986
Average			-2.133	1.128	N/A

Validations were performed to verify the effectiveness of the above model for determining h_{GLR}. We found that the actual in-control ANSS value for h_{GLR} obtained from Equation (4.1) is not always close to the desired ANSS value. Define an error rate as the absolute difference between the actual and the desired ANSS values divided by the desired ANSS value, i.e.

$$\text{Error Rate} = \frac{|\text{Actual ANSS} - \text{Desired ANSS}|}{\text{Desired ANSS}}.$$

We will evaluated the error rate for h_{GLR} obtained from Equation (4.1) relative to the error rate of the Shewhart chart control limit h_S obtained from using either the exact binomial distribution or the normal approximation to the binomial distribution.

Table 4.7 gives error rates for the binomial GLR and Shewhart charts for seven p_0 levels ranging from 0.0001 to 0.2 and two or three sample sizes at each level of p_0. In addition, four in-control ANSS values are specified to represent different in-control performance requirements. As an example, consider the situation when $p_0 = 0.05$ and $n = 100$, with the desired in-control ANSS value being 500. For the Shewhart chart, we first try the normal approximation approach, since this is what has traditionally been used. The normal quantile corresponding to the probability of $1/500 = 0.002$ is 2.878, and this gives $h'_{\mathrm{S}} = 5 + 2.878\sqrt{100 \times 0.05 \times 0.95} = 11.27$. The chart signals if $T_k \geq h'_{\mathrm{S}}$, so we can take h'_{S} to be 12 which gives an ANSS value of 233.96. If the exact binomial distribution is used we can get closer to the desired in-control ANSS of 500 by taking h'_{S} to be 13, and then the in-control ANSS is 682.90. The error rates for

the two cases are 0.53 and 0.37, respectively. For the binomial GLR chart, $h_{\mathrm{GLR}} = 4.88$ is obtained based on Equation (4.1), and the corresponding ANSS value is 394.66, which gives an error rate of 0.21.

TABLE 4.7. Error rates of the in-control ANSS value given by estimated control limit relative to the desired in-control ANSS value, for both Shewhart and binomial GLR charts.

		Desired In-Control ANSS											
		50			500			1000			2000		
		Shewhart			Shewhart			Shewhart			Shewhart		
p_0	n	Binomial	Normal	GLR	Binomial	Normal	GLR	Binomial	Normal	GLR	Binomial	Normal	GLR
0.0001	500	0.59	0.59	1.34	0.66	0.96	0.17	0.17	0.98	0.84	0.59	0.99	0.69
	5000	0.39	0.78	0.11	0.14	0.86	0.10	0.43	0.93	0.01	0.71	0.97	0.06
0.0010	20	0.01	0.01	0.01	0.90	0.90	0.87	0.95	0.95	0.46	0.97	0.97	0.61
	200	0.15	0.15	0.09	0.76	0.89	0.09	0.12	0.94	0.28	0.56	0.97	0.27
	1000	0.06	0.06	0.24	0.45	0.89	0.14	0.70	0.73	0.22	0.15	0.86	0.11
0.0050	20	0.79	1.01	0.47	0.55	0.80	0.23	0.78	0.90	0.36	0.89	0.95	0.19
	200	0.07	0.79	0.23	0.44	0.57	0.12	0.72	0.79	0.19	0.11	0.89	0.09
0.0100	10	0.79	0.79	0.47	0.53	0.45	0.26	0.77	0.28	0.33	0.88	0.64	0.27
	100	0.09	0.09	0.22	0.42	0.89	0.09	0.71	0.71	0.16	0.06	0.85	0.05
	200	0.25	0.61	0.26	0.53	0.53	0.16	0.01	0.77	0.06	0.51	0.88	0.15
0.0500	10	0.74	0.77	0.02	0.83	0.83	0.05	0.03	0.91	0.05	0.51	0.96	0.15
	100	0.29	0.29	0.16	0.37	0.53	0.21	0.32	0.77	0.30	0.08	0.66	0.24
0.1000	5	0.75	0.75	0.12	0.77	0.77	0.14	0.88	0.88	0.16	0.09	0.94	0.42
	10	0.56	0.72	0.04	0.22	0.84	0.30	0.39	0.92	0.13	0.69	0.69	0.05
	50	0.18	0.18	0.18	0.38	0.38	0.26	0.00	0.69	0.22	0.50	0.84	0.16
0.2000	5	0.65	0.65	0.06	0.70	0.70	0.09	0.85	0.85	0.37	0.56	0.93	0.06
	50	0.35	0.35	0.28	0.20	0.20	0.27	0.07	0.60	0.23	0.46	0.46	0.24
Average		0.40	0.51	0.25	0.52	0.71	0.21	0.46	0.80	0.26	0.49	0.85	0.22

We see from Table 4.7 that the in-control ANSS of the binomial GLR chart given by h_{GLR} that is obtained from Equation (4.1) is reasonably close to the desired value if the product of n and p_0 is not too small, say, $n \times p_0$ not smaller than 0.1. A larger sample size tends to reduce the error rate unless p_0 is extremely large, such as $p_0 > 0.1$. For example, when $p_0 = 0.0001$ and $n = 500$, we have $n \times p_0 = 0.05$, and the error rate for the binomial GLR chart is high in most cases. For the combinations of p_0 and n considered here, the average error rate of the binomial GLR chart is much lower than that of the Shewhart chart, no matter if h_{S} is based

on the exact or approximate distribution. Also notice that the Shewhart chart based on the normal approximation tends to have the highest error rate for most of the cases.

Comparisons between the error rates obtained from binomial GLR chart and Shewhart chart are also displayed graphically in Figures 4.3 through 4.9. Each graph represents one level of p_0, ordered from the smallest to the largest. The same conclusions could be drawn from these graphs: the error rate is generally lower for the binomial GLR chart than the Shewhart chart, except for some extreme cases when $n \times p_0$ are too small ($n \times p_0 < 0.1$). For example, in Figure 4.3, we have $p_0 = 0.0001$, $n = 50$, and $n \times p_0 = 0.05$. As a result, the error rate given by the estimated binomial GLR control limit is quite high.

Although a more accurate way to choose h_{GLR} is to run a simulation program, this would probably not be convenient for most practitioners. Thus, for practical applications involving $n \times p_0$ not smaller than 0.1, we recommend that practitioners use Equation (4.1) to obtain h_{GLR}.

FIGURE 4.3. Plots of the error rates of the Shewhart and binomial GLR charts relative to a desired in-control ANSS value at $p_0 = 0.0001$.

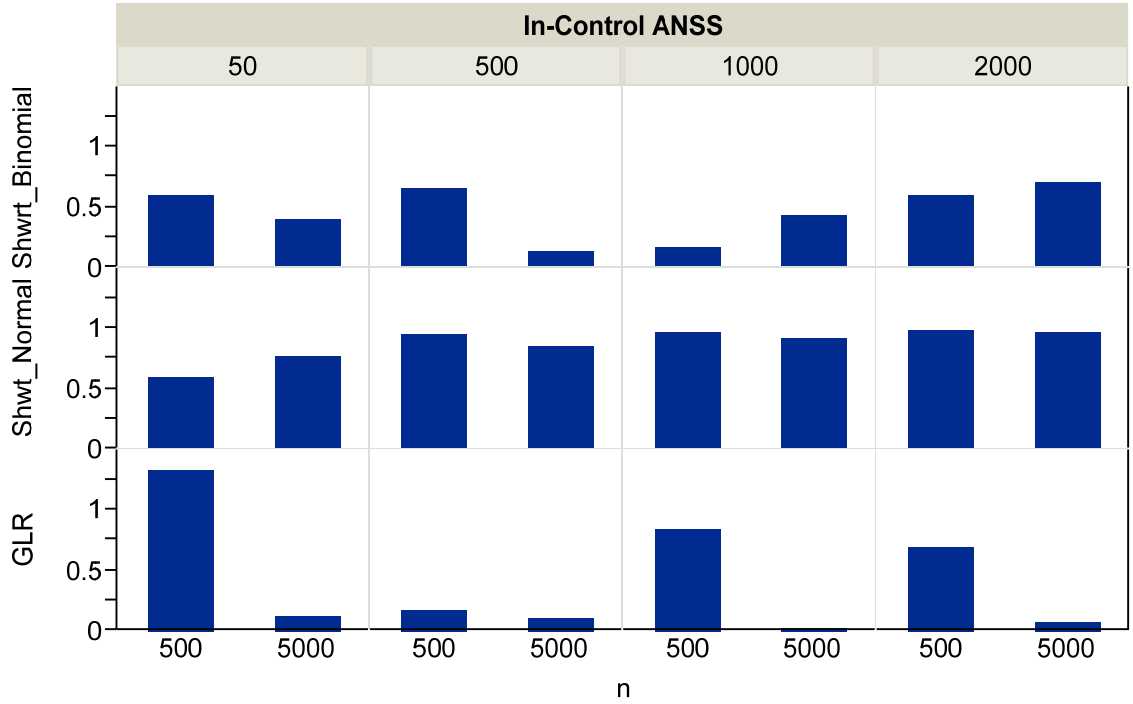


FIGURE 4.4. Plots the of error rates of the Shewhart and binomial GLR charts relative to a desired in-control ANSS value at $p_0 = 0.001$.

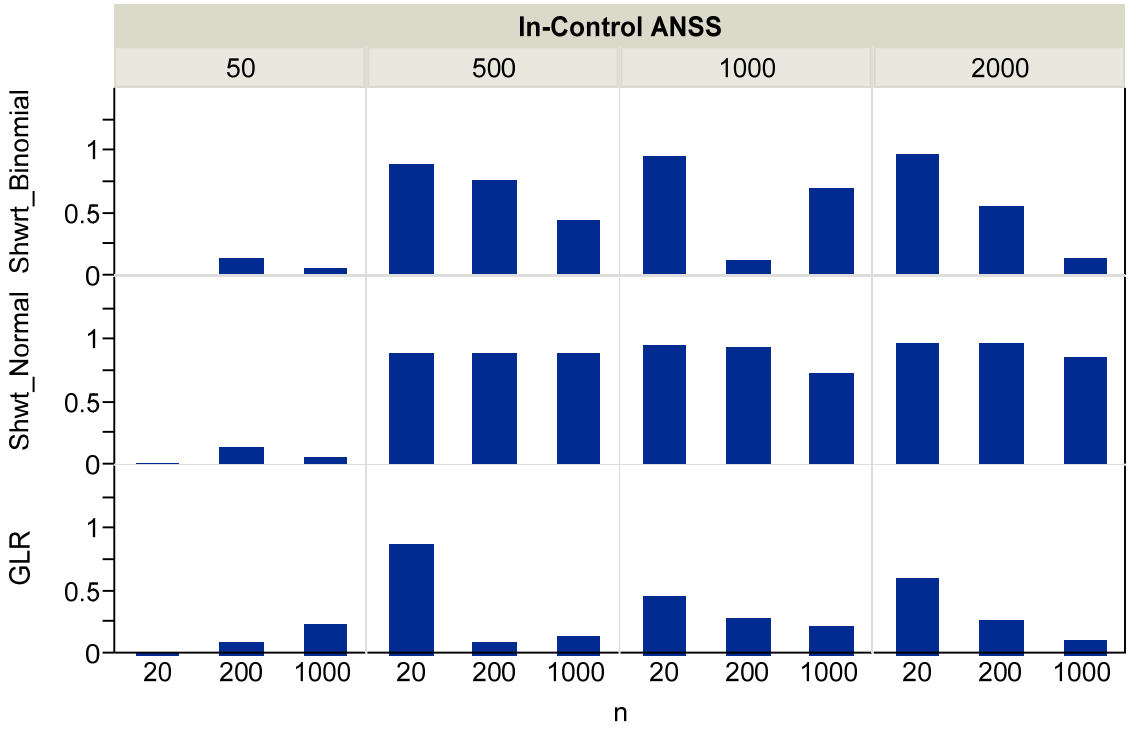


FIGURE 4.5. Plots of the error rates of the Shewhart and binomial GLR charts relative to a desired in-control ANSS value at $p_0 = 0.005$.

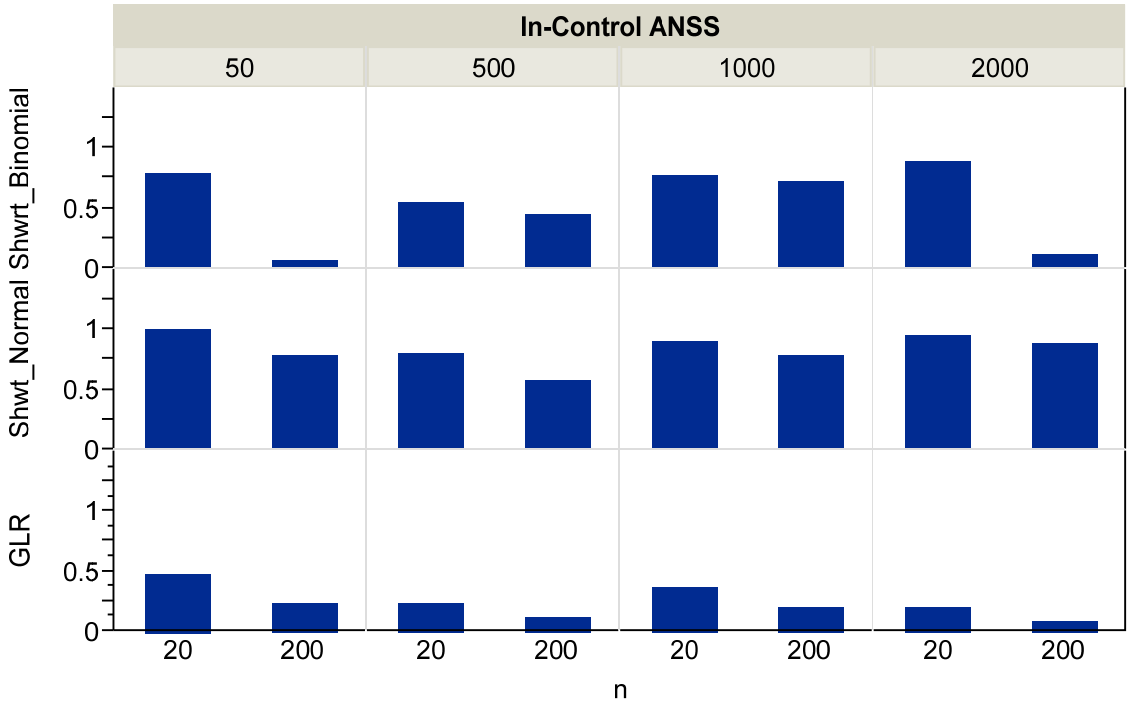


FIGURE 4.6. Plots of the error rates of the Shewhart and binomial GLR charts relative to a desired in-control ANSS value at $p_0 = 0.01$.

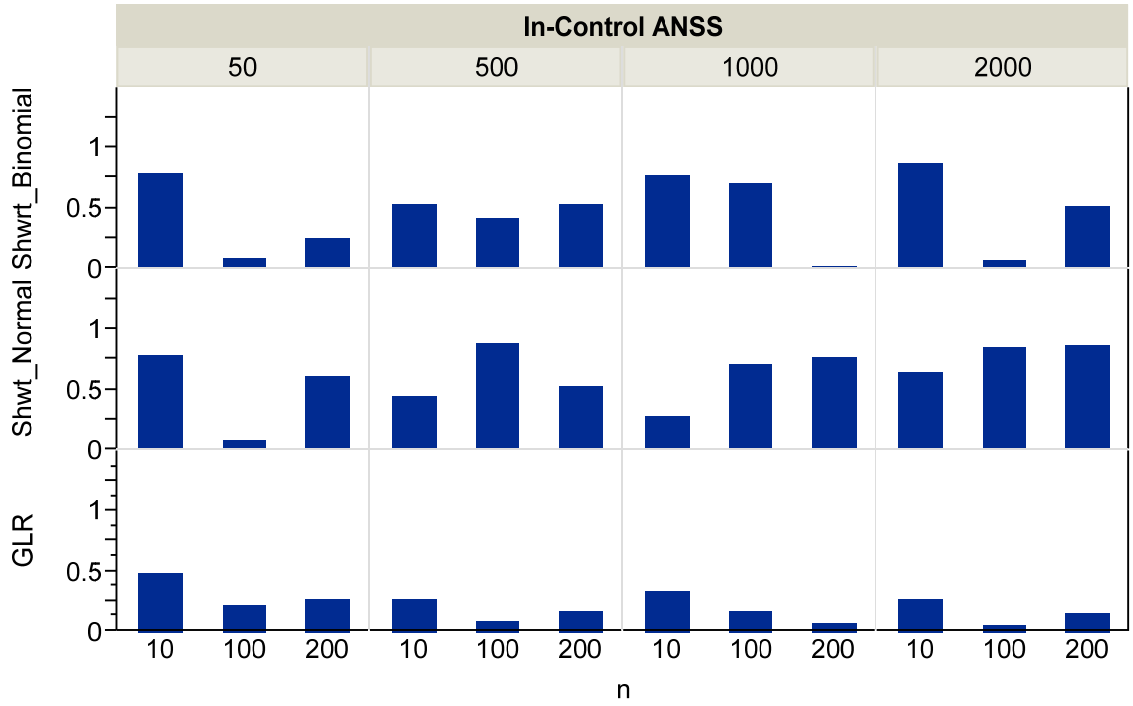


FIGURE 4.7. Plots of the error rates of the Shewhart and binomial GLR charts relative to a desired in-control ANSS value at $p_0 = 0.05$.

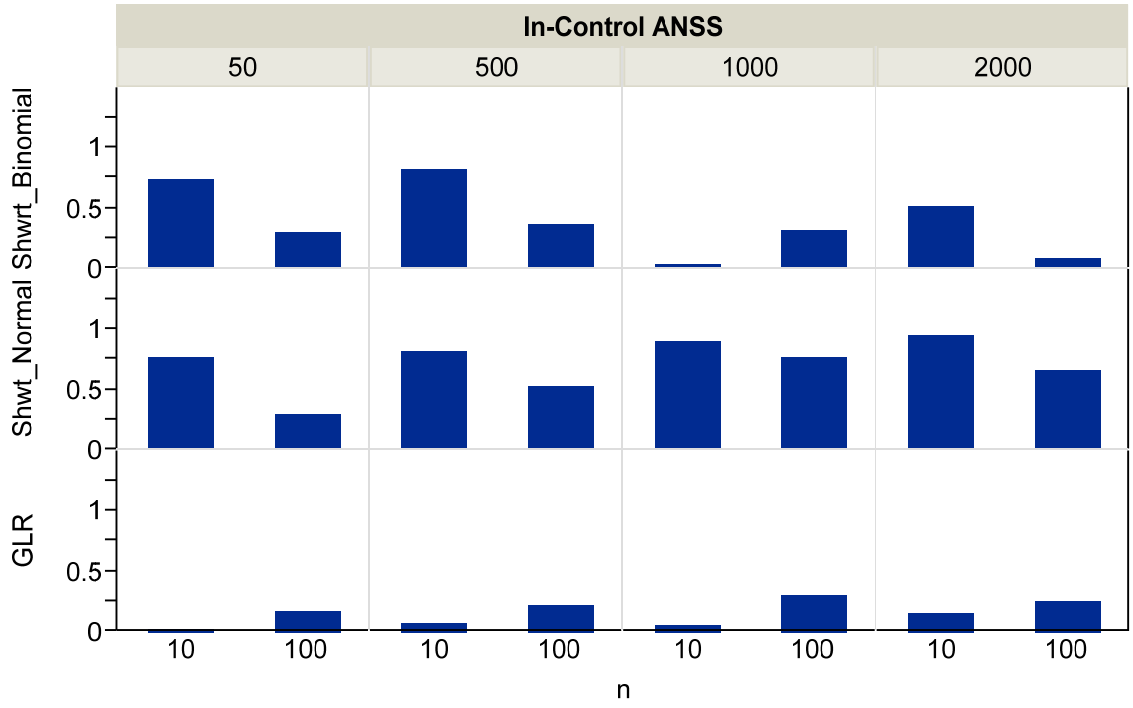


FIGURE 4.8. Plots of the error rates of the Shewhart and binomial GLR charts relative to a desired in-control ANSS value at $p_0 = 0.1$.

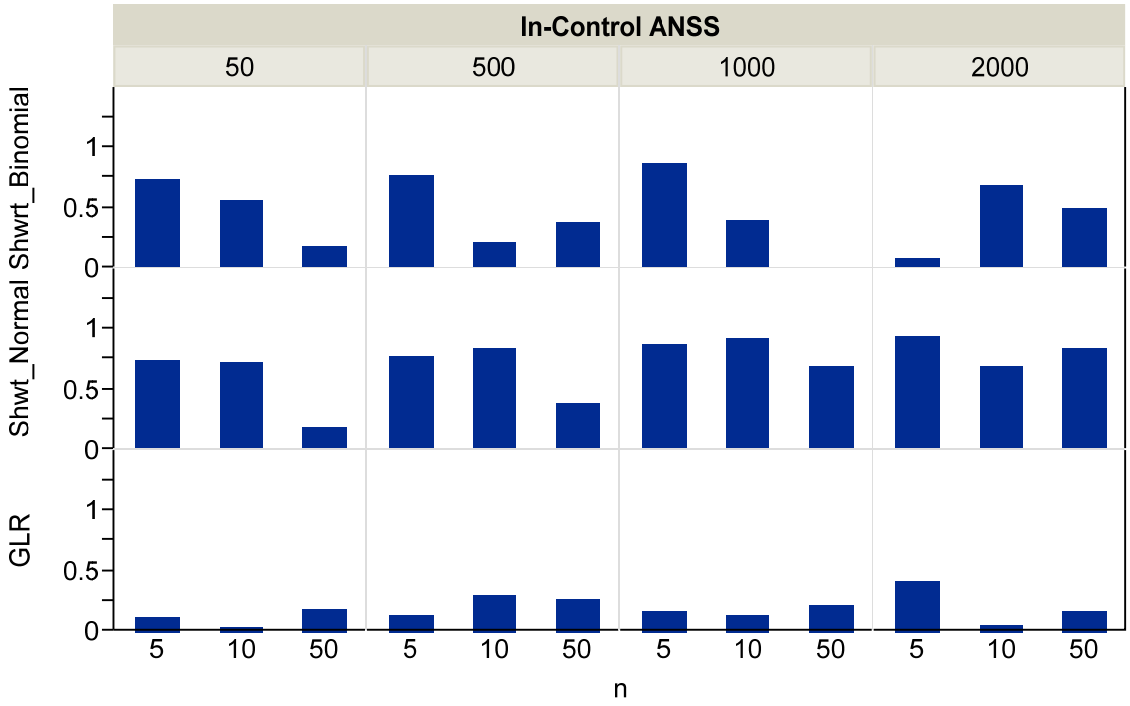
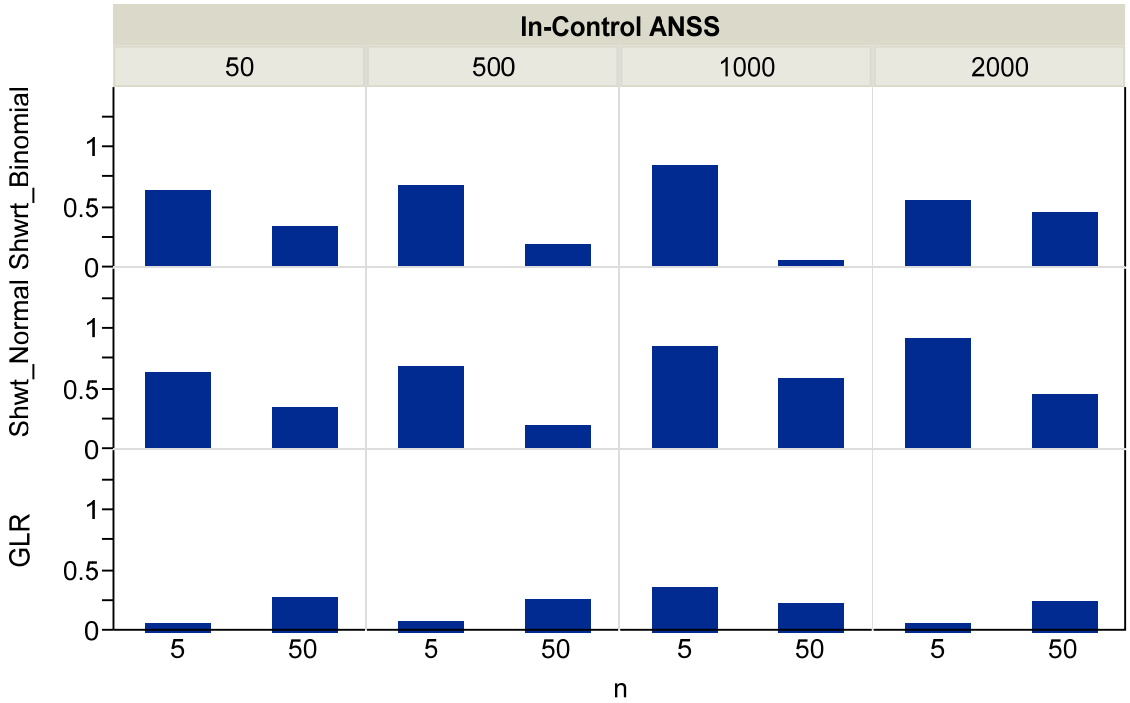


FIGURE 4.9. Plots of the error rates of the Shewhart and binomial GLR charts relative to a desired in-control ANSS value at $p_0 = 0.2$.



Chapter 5

Binomial Control Chart Performance Comparisons

5.1 Comparing the Binomial GLR Chart to the Shewhart np-Chart and the Binomial CUSUM Chart

In this section, we evaluate the performance of the binomial GLR chart relative to the traditional Shewhart np-chart and the binomial CUSUM chart. All of these charts are set up to have approximately the same in-control performance. Comparisons are made across a wide range of shift sizes for p when three levels of p_0 are considered. For each level of p_0, we considered two sample sizes – one is small and one is relatively large. In addition, more investigation is performed for finding the effect of different sample sizes on the performance of the GLR chart when p_0 is relatively small.

For the first comparison, consider the case of $p_0 = 0.01$, $n = 100$, $m = 300$, and $\tau = 100$. In Table 5.1, the columns labeled [2]-[7] give SSANOS values of the Shewhart np-chart, the binomial GLR chart, and four binomial CUSUM charts with different tuning parameters. Columns [8]-[10] are the SSANOS values for three combinations of multiple CUSUM charts, which will be explained in more detail later. Unlike in Tables 4.1 and 4.2 where we reported SSANSS values, in Table 5.1 and the following tables SSANOS values are given because the SSANOS allows for comparisons across different sample sizes. The Shewhart np-chart in column [2] has $h'_S = 5$ which gives an in-control ANSS of 291.35 and an in-control ANOS value of 29,135. The values of h_{GLR} and h_{BinC} were then selected to closely match the in-control performance of the Shewhart chart. The values of p_1 for the four CUSUM charts in columns [4]-[7] range from 0.015 to 0.07.

We see from Table 5.1 that the Shewhart chart only works well for relatively large shift sizes ($p \geq 0.05$). The CUSUM charts are best at the shift sizes close to p_1, and their signal times

might be much longer at other shifts. For instance, the CUSUM chart with $p_1 = 0.015$ works well for small shifts ranging from about 0.013 to 0.017, but not as well for those shifts greater than 0.04; on the other hand, the CUSUM chart with a larger tuning parameter of $p_1 = 0.07$ has quick detection for shifts greater than 0.06, but its performance is not satisfactory for shifts smaller than 0.03. For large shifts such as $p \geq 0.2$, with high probability all charts signal after only one sample. Although the GLR chart cannot beat an individual CUSUM chart at its tuned shift size, the performance of the GLR chart is always close to the best no matter which shift actually occurs. Thus we conclude that the GLR chart has much better over-all performance than the Shewhart and CUSUM charts for detecting multiple shift sizes.

TABLE 5.1. SSANOS values for various control charts with $p_0 = 0.01$ and $n = 100$.

		GLR Chart	Binomial CUSUM Chart				Multiple CUSUM Charts		
n=100	Shewhart Chart	*m*=300	p_1=0.015	p_1=0.020	p_1=0.030	p_1=0.070	4-CUSUM	20-CUSUM	1000-CUSUM
p	*h*=4	*h*=4.13	*h*=2.96	*h*=3.67	*h*=3.94	*h*=3.78	*h*=4.00	*h*=4.13	*h*=4.13
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.010	29134.80	29350.00	29662.13	29301.75	28280.18	29077.20	29448.77	29385.42	29334.60
0.013	9904.40	4677.99	4076.77	4993.56	6322.33	9876.76	4789.45	4700.94	4681.74
0.015	5652.60	2357.89	2095.74	2415.15	3115.92	5617.48	2368.83	2364.95	2357.67
0.017	3524.50	1474.67	1367.13	1457.03	1813.42	3486.06	1473.58	1475.09	1473.45
0.020	1966.10	900.65	893.54	882.10	1003.17	1930.05	902.41	902.75	901.39
0.025	940.71	528.66	570.44	526.22	538.07	910.34	528.37	528.52	528.00
0.030	549.21	372.36	424.36	381.08	365.65	529.24	372.30	372.21	372.63
0.040	269.26	238.58	287.40	255.20	232.35	262.13	238.76	238.87	238.81
0.050	177.21	180.09	223.36	197.45	177.08	175.09	179.96	180.03	180.00
0.060	138.23	148.03	186.53	163.44	146.76	137.63	147.91	148.02	148.00
0.080	109.95	116.82	144.67	126.11	116.54	109.90	116.87	116.82	116.80
0.100	102.39	105.11	121.16	109.43	105.04	102.47	105.10	105.11	105.13
0.150	100.04	100.14	101.51	100.36	100.14	100.04	100.13	100.14	100.13
0.200	100.00	100.00	100.04	100.01	100.00	100.00	100.00	100.00	100.00

Table 5.2 gives SSANOS values for the Shewhart, GLR, and CUSUM charts for the case of $p_0 = 0.01$, $n = 16$, $m = 2{,}000$, and $\tau = 1{,}000$. The Shewhart chart has $h'_S = 3$, which gives an in-control ANOS value of 31,492, and h_{GLR} and h_{BinC} were selected to match the Shewhart chart. The values of p_1 for the CUSUM charts are slightly different from the case of $n = 100$ in Table 5.1, mainly due to the discreteness of the binomial distribution that prevents the desired in-control ANOS value being achieved for some particular value of p_1. From Table 5.2 we find that the Shewhart chart tends to be better than the other charts only for extremely large shift sizes

($p \geq 0.50$); the CUSUM charts are best around their pre-specified shift sizes, but may not perform as well elsewhere; the GLR chart is only slightly worse than the best chart at each shift size but has better overall performance if all of the shift sizes are considered.

The in-control ANOS values of the charts in Table 5.1 are reasonably close to the values in Table 5.2, so this allows for rough comparisons between the case of $n = 100$ and the case of $n = 16$ ($p_0 = 0.01$ in both cases). We see that using $n = 100$ gives a smaller SSANOS when detecting small shifts, while using $n = 16$ gives a smaller SSANOS when detecting large shifts. The effect of different values of n on the charts' performance is further investigated in Section 5.2 for the case when $p_0 = 0.001$ and five different samples sizes are considered.

TABLE 5.2. SSANOS values for various control charts with $p_0 = 0.01$ and $n = 16$.

		GLR Chart	Binomial CUSUM Chart				Multiple CUSUM Charts		
n=16	Shewhart Chart	*m*=2000	p_1=0.015	p_1=0.035	p_1=0.075	p_1=0.200	4-CUSUM	20-CUSUM	1000-CUSUM
p	*h*=2	*h*=6.29	*h*=3.33	*h*=5.25	*h*=5.96	*h*=6.19	*h*=6.21	*h*=6.29	*h*=6.29
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.010	31499.63	31418.39	31664.00	31361.60	30297.60	31464.35	22504.25	31751.75	31346.21
0.013	14764.00	6299.33	4153.99	7407.80	10357.89	14682.75	6138.54	6440.07	6304.60
0.015	9800.40	3075.53	2105.06	3654.61	5862.91	9733.28	3113.73	3116.35	3070.26
0.017	6865.40	1844.69	1358.89	2081.33	3609.91	6763.04	1932.19	1869.00	1846.21
0.020	4341.10	1067.05	875.81	1109.85	1966.54	4248.52	1141.41	1076.41	1066.81
0.025	2334.40	575.32	549.28	550.81	904.63	2246.13	605.62	579.05	574.97
0.030	1418.10	375.39	399.42	350.75	513.64	1336.45	386.78	377.37	375.13
0.040	659.80	211.21	260.59	200.77	241.79	590.94	215.14	212.01	211.38
0.050	372.62	143.78	193.94	141.84	150.18	320.24	145.89	143.95	143.92
0.060	237.79	108.28	155.18	110.59	108.04	199.23	109.23	108.49	108.27
0.080	121.98	72.10	111.85	77.83	69.63	102.27	71.49	72.11	71.99
0.100	75.91	54.18	88.22	60.88	52.34	66.08	52.73	54.19	54.16
0.150	36.47	34.48	58.80	40.29	34.07	34.58	32.12	34.50	34.51
0.200	24.67	26.54	45.08	31.04	26.43	24.33	23.75	26.56	26.54
0.300	17.77	19.57	32.17	22.31	19.54	17.76	17.72	19.57	19.58
0.400	16.30	16.90	25.85	18.13	16.89	16.30	16.29	16.91	16.91
0.500	16.03	16.14	21.42	16.47	16.15	16.03	16.03	16.14	16.15
0.600	16.00	16.01	18.15	16.06	16.01	16.00	16.00	16.01	16.01
0.800	16.00	16.00	16.04	16.00	16.00	16.00	16.00	16.00	16.00

TABLE 5.3. SSANOS values for various control charts with $p_0 = 0.1$ and $n = 97$.

		GLR Chart	Binomial CUSUM Chart				Multiple CUSUM Charts		
n=97	Shewhart Chart	*m*=60	p_1=0.12	p_1=0.14	p_1=0.16	p_1=0.20	4-CUSUM	20-CUSUM	1000-CUSUM
p	*h*=16	*h*=2.81	*h*=2.06	*h*=2.47	*h*=2.47	*h*=2.36	*h*=2.79	*h*=2.81	*h*=2.81
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.100	6233.30	6238.07	6112.15	6380.72	6479.36	6117.85	6397.86	6239.19	6239.16
0.110	2740.20	1891.86	1689.83	2024.06	2308.44	2650.13	1957.15	1891.87	1887.90
0.115	1920.70	1228.83	1102.15	1310.86	1516.71	1847.81	1264.85	1230.79	1227.78
0.120	1390.40	870.54	789.80	910.98	1051.17	1326.53	888.20	871.54	870.82
0.125	1036.40	657.02	606.42	674.37	767.50	983.28	667.80	656.71	656.90
0.130	793.32	520.50	489.48	525.43	585.56	750.58	527.53	520.06	520.03
0.140	501.21	359.60	354.07	358.41	382.12	472.05	363.72	359.91	359.53
0.150	343.68	272.32	279.66	271.32	278.70	325.00	274.52	272.52	272.53
0.175	179.98	171.46	189.54	174.58	169.98	174.63	171.88	171.28	171.43
0.200	126.73	129.60	148.27	134.52	128.67	125.57	129.73	129.76	129.63
0.225	107.18	110.08	124.07	113.82	109.88	106.90	110.13	110.14	110.13
0.250	100.07	101.54	109.53	103.54	101.53	100.03	101.54	101.53	101.53
0.300	97.17	97.31	98.53	97.53	97.31	97.16	97.31	97.31	97.32
0.400	97.00	97.00	97.00	97.00	97.00	97.00	97.00	97.00	97.00

TABLE 5.4. SSANOS values for various control charts with $p_0 = 0.1$ and $n = 10$.

		GLR Chart	Binomial CUSUM Chart				Multiple CUSUM Charts		
n=10	Shewhart Chart	*m*=600	p_1=0.12	p_1=0.18	p_1=0.25	p_1=0.60	4-CUSUM	20-CUSUM	1000-CUSUM
p	*h*=4	*h*=5.11	*h*=2.53	*h*=4.24	*h*=4.79	*h*=4.89	*h*=4.91	*h*=5.11	*h*=5.11
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.100	6116.44	6116.40	6116.60	6105.06	6092.78	6103.68	6485.90	6540.82	6460.24
0.110	3972.90	2407.43	1650.21	2395.44	2905.25	3979.11	2481.77	2449.75	2415.10
0.115	3254.20	1562.09	1069.24	1610.49	2087.82	3239.53	1615.95	1591.27	1568.12
0.120	2691.20	1085.28	757.39	1130.65	1537.71	2678.53	1124.66	1100.97	1089.39
0.125	2244.60	793.70	576.94	823.32	1158.57	2236.32	830.74	809.33	800.26
0.130	1887.90	609.50	461.24	623.90	885.59	1877.99	640.61	620.53	614.52
0.140	1364.90	395.21	326.12	390.35	554.16	1362.02	416.41	402.15	398.70
0.150	1012.80	281.94	251.18	269.08	369.73	1005.20	292.86	285.71	283.15
0.175	527.11	151.94	159.79	142.45	172.80	517.40	153.77	154.28	153.12
0.200	304.82	99.31	117.28	95.25	103.31	295.51	99.46	100.28	99.92
0.250	127.99	56.19	77.52	57.34	54.81	120.28	56.16	56.36	56.35
0.300	66.50	38.71	58.14	41.43	37.60	61.73	38.53	38.58	38.62
0.400	27.28	23.94	39.32	27.05	24.01	25.87	23.57	23.59	23.58
0.500	16.05	17.54	30.32	20.65	18.39	15.81	17.13	17.15	17.15
0.600	12.00	13.85	25.04	17.16	15.12	11.97	13.55	13.56	13.55
0.700	10.49	11.52	21.43	14.60	12.64	10.49	11.37	11.37	11.37
0.800	10.06	10.32	18.88	12.18	10.85	10.07	10.29	10.29	10.29
0.900	10.00	10.02	17.54	10.44	10.09	10.00	10.01	10.01	10.01
1.000	10.00	10.00	16.70	10.00	10.00	10.00	10.00	10.00	10.00

For another example, consider monitoring a process with a higher in-control nonconforming rate of $p_0 = 0.1$. Table 5.3 gives SSANOS values for the case of $n = 97$, $m = 60$, and $\tau = 50$, and Table 5.4 gives SSANOS values for $n = 10$, $m = 600$, and $\tau = 300$. The in-control ANOS for both sample sizes is set at a little over 6,000. Control limits for all the charts are selected in the same way as in the previous examples. The results show that the GLR chart has fairly good performance at all shift sizes; the Shewhart chart does not perform well for small and intermediate shifts; the performance of the four CUSUM charts is more satisfactory around their tuned shift sizes.

For the third example, consider the monitoring of a high quality process with $p_0 = 0.001$. Two sample sizes, $n = 1{,}000$ and 150, are considered, and the corresponding SSANOS results are reported in Table 5.5 and Table 5.6, respectively. The in-control ANOS for both sample sizes is set at around 300,000, which leads to selections of the window sizes for the two GLR charts to be 2,000 and 300, and values of τ to be 100 and 1,000, respectively. Again, we apply the same rules as before to choose the control limits for the GLR and CUSUM charts. The conclusions from Tables 5.5 and 5.6 are the same as those in the previous examples.

TABLE 5.5. SSANOS values for various control charts with $p_0 = 0.001$ and $n = 1{,}000$.

		GLR Chart	Binomial CUSUM Chart				Multiple CUSUM Charts		
n=1000	Shewhart Chart	*m*=300	p_1=0.0012	p_1=0.0020	p_1=0.0035	p_1=0.0070	4-CUSUM	20-CUSUM	1000-CUSUM
p	*h*=4	*h*=4.06	*h*=1.84	*h*=3.55	*h*=3.79	*h*=3.57	*h*=3.79	*h*=4.06	*h*=4.06
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.0010	274961.10	279999.45	275000.08	274994.20	274107.10	276660.40	306619.25	281120.51	280423.56
0.0015	54108.00	23344.14	20242.14	23505.47	31569.15	53617.65	24666.23	23406.25	23397.58
0.0020	19062.00	8983.66	9657.31	8707.12	10276.48	18662.37	9138.67	8978.72	8967.11
0.0025	9210.10	5288.88	6415.87	5246.55	5447.51	8913.36	5399.92	5289.05	5277.35
0.0030	5418.20	3730.32	4863.12	3803.06	3685.66	5228.95	3814.79	3728.79	3731.57
0.0040	2694.70	2395.97	3366.24	2556.99	2329.30	2631.55	2431.40	2394.85	2393.79
0.0050	1786.30	1811.40	2632.69	1979.75	1781.78	1766.28	1820.99	1809.43	1808.04
0.0060	1397.40	1489.67	2198.73	1641.99	1476.78	1393.94	1495.61	1489.75	1490.48
0.0070	1207.50	1298.98	1912.20	1420.11	1292.89	1206.96	1301.69	1298.52	1299.02
0.0080	1109.60	1178.74	1706.52	1272.60	1176.06	1108.98	1180.84	1179.03	1178.90
0.0090	1057.90	1104.18	1552.55	1170.97	1103.46	1057.06	1106.14	1104.67	1104.72
0.0100	1029.70	1059.44	1427.01	1104.87	1058.27	1029.10	1060.13	1059.46	1059.39
0.0150	1000.80	1002.33	1081.48	1005.76	1002.23	1000.83	1002.35	1002.36	1002.37
0.0200	1000.00	1000.06	1007.87	1000.16	1000.06	1000.01	1000.04	1000.05	1000.06
0.0500	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00

TABLE 5.6. SSANOS values for various control charts with $p_0 = 0.001$ and $n = 150$.

		GLR Chart	Binomial CUSUM Chart				Multiple CUSUM Charts		
n=150	Shewhart Chart	*m*=2000	p_1=0.0012	p_1=0.0030	p_1=0.0070	p_1=0.0300	4-CUSUM	20-CUSUM	1000-CUSUM
p	*h*=2	*h*=6.28	*h*=2.05	*h*=5.00	*h*=5.66	*h*=5.87	*h*=6.00	*h*=6.28	*h*=6.28
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.0010	303686.69	303684.09	303186.00	302280.00	303168.00	302711.00	278899.08	307827.88	303533.75
0.0015	95044.00	30672.00	20733.33	32489.88	54856.56	94530.90	35336.74	31219.64	30665.74
0.0020	42335.00	10705.08	9687.95	10114.69	17842.96	41656.12	11642.45	10820.89	10708.88
0.0025	22887.00	5797.06	6321.56	5248.76	8237.72	22369.56	5885.41	5837.37	5792.78
0.0030	13982.00	3781.91	4701.33	3445.97	4736.38	13524.59	3764.54	3811.34	3780.55
0.0040	6565.90	2128.87	3123.44	2030.20	2303.31	6180.92	2111.55	2141.01	2127.82
0.0050	3741.90	1448.34	2348.46	1445.97	1470.52	3456.60	1430.40	1451.24	1445.14
0.0060	2407.30	1087.09	1889.07	1129.49	1075.95	2179.68	1067.93	1087.00	1086.57
0.0070	1684.70	868.94	1584.38	932.44	851.26	1516.37	849.77	870.43	868.37
0.0080	1251.20	725.21	1366.19	800.70	705.87	1122.11	706.66	725.86	725.97
0.0090	975.63	622.89	1202.10	702.94	607.06	875.87	606.53	623.11	622.48
0.0100	787.84	546.57	1076.91	629.60	533.96	713.33	533.21	547.71	547.84
0.0125	518.12	423.77	858.23	503.93	416.15	480.52	414.52	424.26	423.80
0.0150	383.72	350.56	717.98	424.79	346.24	364.49	345.27	350.53	350.58
0.0200	258.89	269.04	546.95	328.64	267.92	254.29	267.92	269.30	269.00
0.0500	152.77	157.21	257.69	166.15	157.13	152.72	157.12	157.21	157.22
0.1000	150.01	150.02	160.03	150.07	150.02	150.00	150.02	150.02	150.02
0.2000	150.00	150.00	150.00	150.00	150.00	150.00	150.00	150.00	150.00

In summary, the results presented in Tables 5.1-5.6 show that the Shewhart chart and individual binomial CUSUM charts are good for certain shift sizes, but if the true shift size is quite different from what is expected, their performance would not be very satisfactory. On the other hand, the binomial GLR chart has better overall performance than the other charts when multiple shift sizes are considered. Therefore, in practical situations, we would recommend using the binomial GLR chart for effective detection of a wide range of process shifts.

5.2 Evaluating Difference Sample Sizes

In the previous section, for a particular p_0, we considered two sample sizes – one large and one relatively small – when we evaluated the performance of the binomial charts. In fact, it is not a trivial issue to choose an appropriate sample size so that the corresponding control chart would have satisfactory performance for monitoring a specific process. Therefore, in this section

we further investigate the effect of different sample sizes on the performance of the binomial GLR chart. A careful examination of the best choice of n requires considering the fact that small samples could presumably be taken more frequently than large samples. We assume that a shift can only occur between samples and that the sampling rate is fixed as one item per unit time such as one minute, in which case it is reasonable to assume that, for example, a sample of $n = 35$ is taken every 35 minutes, a sample of $n = 80$ is taken every 80 minutes, etc. Moreover, suppose that when the n items from the same sample are being inspected, the time between each individual inspection is negligible, so that the total time spent on inspecting one sample is zero.

In addition, instead of using the SSANOS, the steady state average time to signal (SSATS) is used as the performance metric here since it better measures the steady state signal time based on a constant sampling rate. The SSATS assumes that the time of the shift is uniform on the time interval between two consecutive samples. Thus, if the time interval between two samples is d, then the minimum SSATS value would be $d/2$, while the minimum SSANOS always equals to the sample size n.

Simulation is performed based on the high quality process described in the last example in the previous section where we have $p_0 = 0.001$, and the results are given in Table 5.7. In addition to $n = 1{,}000$ and $n = 150$, another three sample sizes ($n = 35$, 80 and 500) are considered here so that a wide range of values of n is covered to allow us to evaluate the effect of using different sample sizes on the control chart performance, and then give a general guideline on how to choose n in practice. The sampling scheme follows the principle that we sample more frequently with a small sample size, and less frequently when the sample size is large. In particular, we assume that the time interval d between two consecutive samples equals to the sample size n, i.e. $d = 35$ when $n = 35$, $d = 80$ when $n = 80$, $d = 150$ when $n = 150$, etc.

TABLE 5.7. SSATS values for various binomial GLR charts with different sample sizes when $p_0 = 0.001$.

	Binomial GLR Charts				
	d=*n*=35 *m*=8600 *τ*=3600	*d*=*n*=80 *m*=3800 *τ*=1575	*d*=*n*=150 *m*=2000 *τ*=840	*d*=*n*=500 *m*=600 *τ*=252	*d*=*n*=1000 *m*=300 *τ*=126
p [1]	*h*=7.37 [2]	*h*=6.68 [3]	*h*=6.28 [4]	*h*=5.05 [5]	*h*=4.06 [6]
0.0010	300012.49	299404.81	303684.09	301084.30	279999.45
0.0015	35590.82	32248.70	30608.75	26155.50	22816.60
0.0020	12142.30	11130.15	10640.84	9340.75	8474.50
0.0025	6423.07	5933.94	5716.49	5130.15	4781.50
0.0030	4116.26	3840.69	3710.90	3385.50	3231.60
0.0040	2238.76	2114.01	2053.16	1919.20	1893.50
0.0050	1470.47	1405.46	1369.44	1307.60	1309.80
0.0060	1069.11	1035.01	1010.97	981.35	989.70
0.0070	829.15	808.88	794.76	782.45	798.40
0.0080	671.01	659.55	650.22	649.80	679.10
0.0090	560.58	554.50	548.55	554.30	604.30
0.0100	479.40	476.22	472.05	484.30	559.40
0.0125	348.15	349.22	348.62	370.15	512.70
0.0150	271.84	272.58	275.64	309.50	502.30
0.0200	187.06	186.42	194.28	262.35	500.10
0.0500	61.75	60.72	82.19	250.00	500.00
0.1000	28.76	40.85	75.02	250.00	500.00
0.2000	18.14	40.00	75.00	250.00	500.00
0.3000	17.51	40.00	75.00	250.00	500.00

In Table 5.7, the change-point τ is set to ensure that the actual change happens after the same number of items is inspected for the binomial GLR charts considered here with five different sample sizes. Notice that for those extremely large shifts ($p \geq 0.3$), the GLR charts always signal after one sample, thus the corresponding SSATS values approach $d/2$ for the steady state. We find that using a small sample such as $n = 35$, the GLR chart gives a smaller SSATS when detecting large shifts, while using $n = 1{,}000$ gives a smaller SSATS when detecting small shifts. In other words, it seems that taking small samples (which could presumably be taken more frequently) would be better for detecting large shifts, while taking larger samples (which would presumably be taken less frequently) would be better for detecting small shifts. Similar conclusions can be drawn for other values of p_0 such as $p_0 = 0.01$ or 0.1.

As a matter of fact, Reynolds and Stoumbous (2000) explored the effect of different samples sizes on the performance of the binomial and Bernoulli CUSUM charts, and the results from their study are similar to what we observe here.

In general, a large n is usually recommended for effectively detecting small shifts and a small n is better for large shifts. However, for the binomial GLR chart to have good performance over a wide range of shift sizes, it probably would be reasonable to consider some intermediate value of n as a compromise. Additionally, in practice the choice of n may depend on convenience. For example, it may be that inspecting a large sample every 4 hours is be more convenient than inspecting a small sample every hour, or the reverse could be true. Therefore, to select a proper sample size n, we need to consider not only factors like the shift size of interest and the value of p_0, but also the issue of practical convenience. On the other hand, although the choice of n may depend on various factors, it generally does not have a very large effect on the performance of the GLR chart unless the shift size is very large, which may not be very likely to occur in many practical applications.

5.3 Comparing the Binomial GLR Chart to the Shewhart-CUSUM Combination

We see from Section 5.1 that the individual binomial CUSUM chart is best around the shift size it is tuned to detect, but is not as good if the actual shift is far from what is expected. Thus its performance over a wide range of shifts is worse than the binomial GLR chart. Instead of using one CUSUM chart to monitor all shifts, one way to improve the overall performance of the binomial CUSUM chart is to use two or more CUSUM charts simultaneously, with each individual chart tuned to detect a particular range of shifts. In the application of monitoring binomial processes, the simplest implement of this idea is to use a Shewhart-np chart together with a binomial CUSUM chart. Since the Shewhart chart is a special case of the CUSUM chart, this Shewhart-CUSUM combination actually is equivalent to using two CUSUM charts together, one for detecting large shifts, and one for detecting small shifts. In this sense, the Shewhart-CUSUM combination is expected to have better overall performance for a wider shift range. In fact, this type of control scheme has been studied extensively and has become a fairly standard approach for monitoring the normal processes over the past decades (see, for example, Lucas

(1982)). However, there seems to be little discussion of applying this method to monitor p based on the binomial distribution.

In this section we evaluate the performance of the binomial GLR chart relative to the combined Shewhart-CUSUM chart as well as the individual binomial CUSUM chart. Simulation is conducted to obtain the ANOS and SSANOS values based on the six p_0 and n combinations being studied previously in Section 5.1, and the results are presented in Tables 5.8 through 5.11. In the Shewhart-CUSUM combination, both the Shewhart and CUSUM charts have its own control limit (denoted by h_S and h_C, respectively), and they are selected so that the in-control ANOS of the combined chart closely matches the in-control ANOS of the Shewhart-np chart, which is the same as how we choose the control limit for the GLR and individual CUSUM charts. However, unlike in the normal case where we can easily find many combinations of h_S and h_C that may be used to achieve the desired in-control performance, for the binomial case, h_S of the Shewhart chart in the combination can only take integer values, which provides us with very limited in-control ANOS values that are achievable. Therefore, in most cases the same control limit used by the standard Shewhart-np chart is used for the Shewhart-np chart in the combination, and we mainly focus on adjusting h_C of the CUSUM chart to obtain the desired in-control ANOS. In addition to that, the tuning parameter p_1 used by the CUSUM chart in the combination is selected in the same way as in the individual CUSUM charts.

For the first example, consider the case where $p_0 = 0.01$. In Table 5.8, columns [3] through [6] give SSANOS values for the binomial GLR chart, two individual binomial CUSUM charts and two Shewhart-CUSUM combinations with $n = 100$, while the results for $n = 16$ are presented similarly in columns [7] through [11]. The value of l for the Shewhart-CUSUM combination is the ratio of the in-control ANOS values of the Shewhart chart and the CUSUM chart, and this value is about 6 for $n = 100$, but larger than 30 for $n = 16$. If a smaller value of h_S is used for $n = 100$ and $n = 16$, say $h_S = 4$ and $h_S = 2$, respectively, then we would have $l < 0.01$ for both cases in order to achieve the desired in-control ANOS values.

We see from Table 5.8 that adding the Shewhart chart to the CUSUM chart only helps significantly when the CUSUM chart is tuned to detect very small shifts. For example, for

$n = 100$, if we compare column [4] to column [3], we see that the Shewhart-CUSUM combination with $p_1 = 0.02$ is better than the CUSUM chart alone with $p_1 = 0.02$ when $p > 0.04$, and is only slightly worse for smaller shifts. On the other hand, when p_1 gets larger, such as $p_1 = 0.03$, the SSANOS values for the individual CUSUM and Shewhart-CUSUM combination in columns [5] and [6] are very similar, in which case there seems to be no need to use the Shewhart chart in addition to the CUSUM chart. We also looked at those cases with $p_1 > 0.03$, and found that the Shewhart-CUSUM combination is no better than the individual CUSUM chart. The same pattern can be found for $n = 16$ when $p_1 = 0.02$, where the SSANOS values of the Shewhart-CUSUM combination is smaller than that of the individual CUSUM chart for large shifts, but slightly larger for small shifts. Their performance is almost the same for all shifts when $p_1 = 0.03$.

Moreover, for $n = 100$, when compared to the binomial GLR chart (column [2]), the Shewhart-CUSUM combination with $p_1 = 0.02$ (column [4]) is only a little bit better at its tuned shift $p = 0.02$, but worse for all other shifts. The Shewhart-CUSUM combination with $p_1 = 0.03$ (column [6]) is not good for detecting small shifts; for shifts around 0.03, its performance is slight better than the GLR chart; for larger shifts such as $p > 0.06$, its performance is almost the same as the GLR chart. When a smaller sample size like $n = 16$ is considered, the Shewhart-CUSUM combination with a small value of p_1 is better than the GLR chart for small shifts. However, adding the Shewhart chart does not improve its performance for detecting large shifts. Using a larger value of p_1 for the Shewhart-CUSUM combination hurts its ability to effectively detect small shifts, yet with no gain for large shifts. Thus, in terms of better overall performance, rather than adding a Shewhart chart to the CUSUM chart, it is better to use a GLR chart, since designing such a combination requires a large amount of extra effort but without gaining much in performance for detecting large shifts.

The results given in Table 5.9 are for $p_0 = 0.1$ and $n = 97$. Five combinations of the Shewhart and CUSUM charts are evaluated, with l ranges from 1.59 to 4.72. Nevertheless, none of those combinations is much better than the individual binomial CUSUM chart with the same value of p_1, even for large shifts. The binomial GLR chart is only slightly worse than the Shewhart-CUSUM combination when p_1 is small for detecting small shifts, and is almost always

better than all the Shewhart-CUSUM combinations for detecting large shifts. Again, the use of the Shewhart chart in conjunction of the CUSUM chart fails to make it more sensitive to the large shifts.

Tables 5.10 considers another sample size $n = 10$ for $p_0 = 0.1$. In addition, the results for a very small in-control proportion $p_0 = 0.001$ and $n = 1{,}000$ or $n = 150$ are given in Table 5.11. In both tables we find the same conclusions as those we draw from Tables 5.8 and 5.9.

In summary, from the comparisons among the binomial GLR chart, the individual binomial CUSUM charts and the Shewhart-CUSUM combinations, we find that when p_1 is small, the CUSUM chart or the Shewhart-CUSUM combination may be better than the GLR chart for small shifts, but worse for the other shifts. When p_1 is large, the CUSUM chart or the Shewhart-CUSUM combination may be slightly better than the GLR chart for large shifts, but much worse for small shifts. In addition, the values of p where one chart may be better than the other also depend on p_0 and n. When comparing the Shewhart-CUSUM combination to the CUSUM chart with the same value of p_1, it seems the Shewhart chart only helps for detecting large shifts if p_1 is small. For large values of p_1, the overall performance of the Shewhart-CUSUM combination is no better than that of the individual CUSUM chart.

Table 5.8. SSANOS values for the binomial GLR, binomial CUSUM, and Shewhart-CUSUM combination charts with $p_0 = 0.01$.

	n=100					*n*=16				
	GLR	CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM	GLR	CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM
	m=300	p_1=0.02	*l*=5.80, p_1=0.02	p_1=0.03	*l*=6.62, p_1=0.03	*m*=2000	p_1=0.02	*l*=30.14, p_1=0.02	p_1=0.035	*l*=31.30, p_1=0.035
p	*h*=4.13	*h*=3.67	h_S=5, h_C=3.68	*h*=3.94	h_S=5, h_C=3.94	*h*=6.29	*h*=4.26	h_S=3,h_C=4.27	*h*=5.25	h_S=3, h_C=5.26
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
0.010	29350.00	29301.75	29923.90	28280.18	28299.30	31418.39	30919.66	31401.34	31361.60	30946.84
0.013	4677.99	4993.56	5064.54	6322.33	6329.72	6299.33	5073.84	5164.93	7407.80	7358.50
0.015	2357.89	2415.15	2434.13	3115.92	3113.79	3075.53	2418.92	2445.68	3654.61	3639.42
0.017	1474.67	1457.03	1475.62	1813.42	1812.05	1844.69	1448.54	1462.63	2081.33	2082.32
0.020	900.65	882.10	888.73	1003.17	1004.78	1067.05	859.92	867.50	1109.85	1107.88
0.025	528.66	526.22	532.10	538.07	538.12	575.32	500.96	504.46	550.81	551.10
0.030	372.36	381.08	382.49	365.65	365.61	375.39	352.82	353.95	350.75	350.63
0.040	238.58	255.20	248.61	232.35	232.46	211.21	222.71	222.53	200.77	200.58
0.050	180.09	197.45	186.39	177.08	177.24	143.78	163.83	162.78	141.84	141.33
0.060	148.03	163.44	151.53	146.76	146.74	108.28	130.13	128.57	110.59	109.98
0.070	128.79	141.21	130.49	128.17	128.13	86.66	108.31	106.35	91.11	90.52
0.080	116.82	126.11	117.68	116.54	116.51	72.10	92.96	90.65	77.83	77.19
0.090	109.56	116.06	109.83	109.31	109.30	61.79	81.71	78.97	68.27	67.30
0.100	105.11	109.43	105.28	105.04	105.03	54.18	73.01	69.89	60.88	59.80
0.150	100.14	100.36	100.15	100.14	100.13	34.48	48.86	43.80	40.29	38.56
0.200	100.00	100.01	100.00	100.00	100.00	26.54	37.89	31.42	31.04	28.66
0.300	100.00	100.00	100.00	100.00	100.00	19.57	27.36	20.54	22.31	19.99
0.400	100.00	100.00	100.00	100.00	100.00	16.90	21.66	17.05	18.13	16.98
0.500	100.00	100.00	100.00	100.00	100.00	16.14	18.12	16.16	16.47	16.16
0.600	100.00	100.00	100.00	100.00	100.00	16.01	16.49	16.01	16.06	16.01

Table 5.9. SSANOS values for the binomial GLR, binomial CUSUM, and Shewhart-CUSUM combination charts with $p_0 = 0.1$ and $n = 97$.

	GLR	CUSUM	Shew-CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM
n=97	m=60	p_1=0.12	l=1.59, p_1=0.12	l=4.49, p_1=0.12	p_1=0.14	l=1.72, p_1=0.14	l=4.72, p_1=0.14	p_1=0.16	l=2.04, p_1=0.16
p	h=2.81	h=2.06	h_S=17, h_C=2.29	h_S=18, h_C=2.13	h=2.47	h_S=17, h_C=2.64	h_S=18, h_C=2.47	h=2.47	h_S=17, h_C=2.48
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.100	6238.07	6112.15	6288.54	6209.97	6380.72	6561.81	6357.47	6479.36	6507.41
0.110	1891.86	1689.83	1811.95	1731.26	2024.06	2138.66	2026.51	2308.44	2312.04
0.115	1228.83	1102.15	1175.21	1125.63	1310.86	1378.21	1311.31	1516.71	1518.88
0.120	870.54	789.80	836.15	804.64	910.98	953.43	909.74	1051.17	1053.78
0.125	657.02	606.42	638.34	615.89	674.37	701.55	674.19	767.50	767.87
0.130	520.50	489.48	512.08	495.75	525.43	544.28	525.79	585.56	585.77
0.140	359.60	354.07	364.22	355.47	358.41	367.35	358.07	382.12	382.36
0.150	272.32	279.66	281.26	278.24	271.32	275.72	271.46	278.70	278.38
0.175	171.46	189.54	177.80	181.74	174.58	171.96	174.62	169.98	169.78
0.200	129.60	148.27	132.25	137.67	134.52	129.84	134.41	128.67	128.75
0.250	101.54	109.53	101.74	103.91	103.54	101.59	103.52	101.53	101.51
0.300	97.31	98.53	97.32	97.57	97.53	97.30	97.54	97.31	97.31
0.400	97.00	97.00	97.00	97.00	97.00	97.00	97.00	97.00	97.00

Table 5.10. SSANOS values for the binomial GLR, binomial CUSUM, and Shewhart-CUSUM combination charts with $p_0 = 0.1$ and $n = 10$.

	GLR	CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM
n=10	m=600	p_1=0.12	l=11.07, p_1=0.12	p_1=0.18	l=11.30, p_1=0.18	p_1=0.25	l=11.18, p_1=0.25
p	h=5.11	h=2.53	h_S=5, h_C=2.59	h=4.24	h_S=5, h_C=4.3	h=4.79	h_S=5, h_C=4.79
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
0.100	6116.40	6116.60	6139.96	6105.06	6026.28	6092.78	5898.04
0.110	2407.43	1650.21	1682.55	2395.44	2425.69	2905.25	2848.06
0.115	1562.09	1069.24	1091.23	1610.49	1634.41	2087.82	2053.39
0.120	1085.28	757.39	772.95	1130.65	1146.89	1537.71	1513.87
0.125	793.70	576.94	587.45	823.32	838.10	1158.57	1141.59
0.130	609.50	461.24	469.70	623.90	632.58	885.59	877.69
0.140	395.21	326.12	331.61	390.35	395.03	554.16	548.10
0.150	281.94	251.18	255.28	269.08	272.56	369.73	368.06
0.175	151.94	159.79	161.43	142.45	143.71	172.80	171.97
0.200	99.31	117.28	117.86	95.25	95.44	103.31	102.85
0.225	56.19	93.17	76.18	71.44	57.17	71.81	54.63
0.300	38.71	58.14	55.48	41.43	41.00	37.60	37.23
0.400	23.94	39.32	33.59	27.05	26.11	24.01	23.21
0.500	17.54	30.32	21.74	20.65	18.66	18.39	17.05
0.600	13.85	25.04	15.01	17.16	14.10	15.12	13.52
0.700	11.52	21.43	11.65	14.60	11.48	12.64	11.35
0.800	10.32	18.88	10.32	12.18	10.75	10.85	10.69
0.900	10.02	17.54	10.02	10.44	10.31	10.09	10.29
1.000	10.00	16.70	10.00	10.00	10.02	10.00	10.01

Table 5.11. SSANOS values for the binomial GLR, binomial CUSUM, and Shewhart-CUSUM combination charts with $p_0 = 0.001$.

	n=1000					*n*=150				
	GLR	CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM	GLR	CUSUM	Shew-CUSUM	CUSUM	Shew-CUSUM
	m=300	p_1=0.002	*l*=5.91, p_1=0.002	p_1=0.0035	*l*=6.23, p_1=0.0035	*m*=2000	p_1=0.002	*l*=26.66, p_1=0.002	p_1=0.003	*l*=26.66, p_1=0.003
p	*h*=4.06	*h*=3.55	h_S=5, h_C=3.63	*h*=3.79	h_S=5, h_C=3.78	*h*=6.28	*h*=4.22	h_S=3, h_C=4.24	*h*=5.00	h_S=3, h_C=5
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
0.0010	279999.45	274994.20	287793.56	274107.10	272712.99	303684.09	304976.53	304218.91	302280.00	304218.91
0.0015	23344.14	23505.47	24208.54	31569.15	31623.05	30672.00	24153.82	24271.12	32489.88	32866.87
0.0020	8983.66	8707.12	8884.18	10276.48	10263.04	10705.08	8624.06	8658.55	10114.69	10188.95
0.0025	5288.88	5246.55	5315.98	5447.51	5448.75	5797.06	5040.34	5049.41	5248.76	5281.18
0.0030	3730.32	3803.06	3827.73	3685.66	3682.65	3781.91	3548.21	3558.66	3445.97	3474.20
0.0040	2395.97	2556.99	2490.38	2329.30	2333.20	2128.87	2240.17	2241.54	2030.20	2051.51
0.0050	1811.40	1979.75	1871.91	1781.78	1779.79	1448.34	1647.97	1640.95	1445.97	1461.25
0.0060	1489.67	1641.99	1526.22	1476.78	1478.65	1087.09	1310.37	1295.66	1129.49	1138.10
0.0070	1298.98	1420.11	1316.81	1292.89	1293.46	868.94	1091.83	1072.80	932.44	937.13
0.0080	1178.74	1272.60	1187.69	1176.06	1175.68	725.21	937.85	914.51	800.70	797.44
0.0090	1104.18	1170.97	1108.82	1103.46	1102.98	622.89	823.37	796.96	702.94	696.14
0.0100	1059.44	1104.87	1061.06	1058.27	1058.37	546.57	735.80	705.63	629.60	618.87
0.0125	1012.63	1026.21	1012.86	1012.49	1012.16	423.77	585.84	546.56	503.93	487.77
0.0150	1002.33	1005.76	1002.38	1002.23	1002.23	350.56	491.22	444.68	424.79	404.06
0.0200	1000.06	1000.16	1000.05	1000.06	1000.06	269.04	380.22	322.89	328.64	302.93
0.0500	1000.00	1000.00	1000.00	1000.00	1000.00	157.21	188.42	158.19	166.15	158.05
0.1000	1000.00	1000.00	1000.00	1000.00	1000.00	150.02	150.46	150.02	150.07	150.02
0.2000	1000.00	1000.00	1000.00	1000.00	1000.00	150.00	150.00	150.00	150.00	150.00

5.4 Combination of Multiple Binomial CUSUM Charts

We know that the binomial CUSUM chart has very good performance at shifts close to the specified p_1, so one way to obtain good performance over a wide range of shifts is to use multiple CUSUM charts together as a combination. Some CUSUM charts in the combination would be tuned to detect small shifts, others to detect intermediate shifts, and others to detect large shifts. Although this idea has been used in monitoring the parameters of a normal process (see, for example, Han et al. (2007), and Reynolds and Lou (2010)), there does not seem to be any previous investigation applying this idea in the case of monitoring p. In the previous section we considered using the Shewhart-np chart together with a binomial CUSUM chart, and found that adding the Shewhart chart to the CUSUM chart does not help much in improving the overall performance of the combined chart. In this section, we will first investigate the idea of using multiple CUSUM charts in more detail and consider the situation where a combination of more than two binomial CUSUM charts is used to monitor p.

To design such a combined chart, first, we need to choose the number of individual binomial CUSUM charts to be used in the combination. Previous researches on using multiple CUSUM charts as a combination to monitor the mean of a normal process show that people tend to use a small number of individual CUSUMs in the combination that ranges from 2 to 5, so here we choose to include four CUSUM charts. Next, the values of p_1 for these four CUSUM charts are selected in a way that a wide range of different shift sizes is covered. Furthermore, in order to simplify the design of such a combination, we use the same h_{BinC} for all the four individual CUSUM charts, in which case individually they have different in-control performances, but as a combination the in-control ANOS value matches that of the corresponding Shewhart chart.

In Tables 5.1-5.6, columns [8] give the SSANOS values for the 4-CUSUM combination under various different situations. Notice that due to the discreteness of the binomial distribution, in Table 5.2 the control limit of the 4-CUSUM combination gives an in-control ANOS value much smaller than all the other charts, thus it is not fair to compare its performance to that of the remaining charts. But in the other tables, we find that by simultaneously applying multiple CUSUM charts as a combination, the performance is improved compared to the individual CUSUM charts in columns [4]-[7] that now it can more effectively detect a wide range of shifts.

Moreover, it seems that the 4-CUSUM combination has very similar performance to the binomial GLR chart, especially for relatively large shifts, and their control limits are also very close. This finding motivates our further research on developing the relationship between the binomial GLR chart and the binomial CUSUM chart. Actually Reynolds and Lou (2010) studied the relationship between the GLR and CUSUM charts for monitoring the normal process mean, and found that the GLR chart without a window was equivalent to an infinite set of CUSUM charts. In addition, their simulation results showed that a set consisting of a few hundreds of CUSUM charts can be used to give a very good approximation to the GLR chart.

To develop the relationship between the binomial GLR chart and the binomial CUSUM charts, first, we see from Equation (3.2) that the binomial GLR statistic without any window is based on the maximum likelihood estimate of p_1, for which all possible values of $p_1 > p_0$ are evaluated. On the other hand, in Equation (2.2), the binomial CUSUM statistic considers a particular p_1 value that is always user-specified rather than estimated. Thus, the GLR chart must be equivalent to an infinite number of CUSUM charts, where all values of $p_1 > p_0$ are included in the set of CUSUM charts and $h_{\mathrm{BinC}} = \mathrm{h}_{\mathrm{GLR}}$ in all of the CUSUM charts.

We now show that GLR chart is actually equivalent to a countable set of CUSUM charts. For a GLR chart to detect an increase in p from p_0, we signal at sample k if for some τ and some $p_1 > p_0$ we have

$$\ln\frac{f(T_{\tau+1}, T_{\tau+2}, \dots, T_k|p_1)}{f(T_{\tau+1}, T_{\tau+2}, \dots, T_k|p_0)} = \ln\frac{p_1^{\sum_{i=\tau+1}^{k} T_i}(1-p_1)^{n(k-\tau)-\sum_{i=\tau+1}^{k} T_i}}{p_0^{\sum_{i=\tau+1}^{k} T_i}(1-p_0)^{n(k-\tau)-\sum_{i=\tau+1}^{k} T_i}}$$

$$= r_2 \sum_{i=\tau+1}^{k} T_i - n(k-\tau)r_1 > h_{\mathrm{GLR}}, \qquad (5.1)$$

where r_1 and r_2 are given by Equation (2.1) and both are functions of p_1. Note that Equation (5.1) is equivalent to

$$\sum_{i=\tau+1}^{k} T_i > \frac{h_{\mathrm{GLR}}+n(k-\tau)r_1}{r_2}. \qquad (5.2)$$

For given values of h_{GLR}, n, and $k-\tau$, the right hand side of Equation (5.2) is minimized with respect to p_1 when

$$p_1 r_2 - r_1 = \frac{h_{\mathrm{GLR}}}{n(k-\tau)}, \tag{5.3}$$

assuming that a $p_1 \in (p_0, 1)$ exists that will satisfy Equation (5.3). If the GLR chart signals for some τ and p_1 such that Equation (5.2) holds, then there must be a solution of p_1 to Equation (5.3) that will also satisfy Equation (5.2).

The possible values of $k - \tau$ are 1, 2, …, k. For a given h_{GLR} and n, there might be no solution to Equation (5.3) for small values of $k - \tau$. Let $s \geq 1$ be the smallest value of $k - \tau$ such that Equation (5.3) has a solution, and denote these solutions by $p_{1,s},\ p_{1,s+1}, \ldots, p_{1,k}$ for $k - \tau = s,\ s+1,\ \ldots, k$. Notice that in Equation (5.3) the left hand side $p_1 r_2 - r_1$ is an increasing function of p_1 and the right hand side $\frac{h_{\mathrm{GLR}}}{n(k-\tau)}$ is a decreasing function of $k - \tau$, thus the sequence of solutions $p_{1,s},\ p_{1,s+1}, \ldots, p_{1,k}$ is a decreasing sequence. It follows that the GLR chart is equivalent to a countable set of CUSUM charts with $h_{\mathrm{BinC}} = h_{\mathrm{GLR}}$ and tuning parameter values of $p_{1,s},\ p_{1,s+1}, \ldots, p_{1,k}$, where $k \to \infty$ because we are not assuming an upper bound on the number of samples that can be taken.

From this we see that the binomial GLR chart without a window is equivalent to a countable set of binomial CUSUM charts. The binomial GLR chart with a window of size m does not have the exact equivalence, but should be approximately equivalent to a finite set of binomial CUSUM charts with $p_1$values fall in the interval of $(p_0, p_{1,s}]$. For each combination of n and p_0, the $p_{1,s}$ value is obtained through numerical analysis, and then used as an upper bound for the $p_1$values for individual CUSUM charts in the CUSUM chart combination. Sometimes adjustments were made if for a particular $p_{1,s}$ value we could not find a control limit that gives the desired in-control performance. Then a value $p'_{1,s} > p_{1,s}$ with satisfactory in-control performance will be chosen, and values of p_1 that fall in the interval of $(p_0, p'_{1,s}]$ will be used.

Simulations were performed to validate the approximate equivalence between the binomial GLR chart with a moving window and a finite set of binomial CUSUM charts. Two examples are given here – both with an in-control value $p_0 = 0.01$, but are for two sample sizes 100 and 16. The same window sizes are used here as before – $m = 300$ for $n = 100$, and $m = 2{,}000$ for $n = 16$. We evaluated six CUSUM combinations that used various numbers of individual CUSUM charts ranging from 20 to 1,000 in this approximation. Notice that rather

than choosing equally spaced values of p_1 within the interval $(p_0, p_{1,s}]$, it seems that values of p_1 need to be selected more closely together when they are close to p_0, and more widely spaced values of p_1 could be used when they are closer to $p_{1,s}$. To implement this idea, a function $p_1(i) = p_0 + p_{1,s}\, i^2/u^2\,, i = 1, 2, \ldots, u$ was used, where u is the total number of CUSUM charts to be used. Also, the range of p_1 could be quite different from one case to another depending on different values of n. Moreover, the same control limit used for the binomial GLR chart was used for both sets of CUSUM charts.

Table 5.12 gives the results for $n = 100$. From this table we find that even if a small number of CUSUM charts, say 20, are used in approximating the GLR chart, the results are quite close to the actual values, especially for shift sizes greater than 0.015. Increasing the number of CUSUM charts used here does not seem to help much in further reducing the differences – there appears to be no substantial improvement from using 50 CUSUM charts to using 1,000 CUSUM charts. On the other hand, for a smaller sample size such as $n = 16$, more CUSUM charts are needed to achieve a relatively accurate approximation. Results displayed in Table 5.13 indicate that the approximations given by using only a few of CUSUM charts are not particularly appealing. Both the in-control and out-of-control performance is not quite precise unless large shift sizes such as $p \geq 0.04$ are considered. The approximation becomes more accurate as the number of CUSUM charts being used increases, and using 200 CUSUM charts seems to be sufficiently large for obtaining a very precise approximation. In fact, a number smaller than 200 would be quite adequate for any approximation needed in practice.

TABLE 5.12. Approximating the binomial GLR chart with a set of binomial CUSUM charts when $p_0 = 0.01,\ n = 100$, and $h_{\mathrm{GLR}} = 4.13$.

		Number of Binomial CUSUM Charts Used					
n=100	GLR	20	50	100	200	500	1000
0.010	29350.00	29381.58	29310.78	29313.94	29311.76	29312.82	29313.54
0.013	4680.56	4685.35	4681.86	4683.74	4682.85	4681.28	4682.29
0.015	2358.74	2365.59	2354.45	2354.41	2354.38	2354.68	2356.00
0.017	1473.98	1473.01	1473.48	1473.77	1473.39	1471.68	1471.81
0.020	904.28	901.93	901.51	903.14	903.93	903.86	903.81
0.025	533.51	532.47	532.17	531.57	531.53	531.48	531.83
0.030	378.19	377.09	376.76	376.75	376.78	376.77	376.75
0.040	246.59	243.94	244.13	244.21	243.91	243.88	244.01
0.050	188.60	185.52	185.53	185.53	185.62	185.46	185.56
0.060	156.54	153.39	153.50	153.44	153.30	153.46	153.23
0.080	122.65	119.67	119.70	119.74	119.72	119.76	119.79
0.100	105.44	102.86	102.79	102.92	102.86	102.91	102.86
0.200	75.65	74.49	74.48	74.44	74.41	74.45	74.46
0.500	59.68	59.21	59.18	59.14	59.13	59.13	59.18
1.000	54.59	54.21	54.35	54.23	54.22	54.22	54.21

TABLE 5.13. Approximating the binomial GLR chart with a set of binomial CUSUM charts when $p_0 = 0.01,\ n = 16$, and $h_{\mathrm{GLR}} = 6.29$.

		Number of Binomial CUSUM Charts Used					
n=16	GLR	20	50	100	200	500	1000
0.010	31418.39	31750.98	31374.26	31372.16	31364.58	31364.95	31367.88
0.013	6302.91	6433.90	6314.04	6305.35	6304.54	6301.47	6298.08
0.015	3077.29	3112.95	3071.58	3069.55	3068.28	3067.76	3067.65
0.017	1846.31	1867.79	1847.94	1846.24	1847.58	1846.30	1846.37
0.020	1064.58	1077.80	1067.53	1067.23	1066.96	1067.34	1067.08
0.025	576.74	579.84	575.53	575.25	575.40	575.40	575.44
0.030	375.51	378.00	375.74	375.57	375.75	375.82	375.70
0.040	212.00	212.49	211.39	211.45	211.45	211.35	211.29
0.050	144.57	144.59	144.19	144.19	144.21	144.24	144.18
0.060	109.01	108.99	108.86	108.98	108.87	108.92	108.94
0.080	73.04	72.82	72.84	72.86	72.81	72.87	72.89
0.100	55.07	54.83	54.78	54.79	54.81	54.80	54.85
0.200	27.46	26.94	26.95	26.96	26.93	26.93	26.97
0.500	15.33	14.90	14.91	14.90	14.91	14.90	14.90
1.000	11.37	11.16	11.16	11.16	11.15	11.15	11.15

Since now we can easily use a set of binomial CUSUM charts to represent the binomial GLR chart, in practice we can also apply a set of CUSUM charts to monitor a process proportion. Using this approach avoids the need to specify the window size of the GLR chart, but it still requires that we use the same control limit that is used by the GLR chart. Moreover, we also need to determine how many CUSUM charts to use in the combination, which introduces another parameter for this approach. The estimate of the shift size $\hat{p}_1$ is given by the p_1 value of the CUSUM chart with the largest value of the CUSUM statistic $C_{\text{Bin},k}$. In terms of plotting the combination of multiple binomial CUSUM charts, although all of the CUSUM statistics would be computed internally by the software, it is not practical to have all of them plotted simultaneously; instead, only the maximum of those CUSUM statistics would actually be plotted. In fact, since all of the individual CUSUM charts use the same control limit h_{GLR}, practitioner would find that a combined CUSUM chart looks very similar to a regular CUSUM or GLR chart (for example plots of regular CUSUM and GLR charts, see Figures 3.2 and 3.3 in Section 3.2), and it would signal if any of the CUSUM statistic goes beyond the control limit.

The simulation results for using either 20 or 1,000 individual binomial CUSUM charts to approximate the binomial GLR chart are presented in columns [9] and [10] in Table 5.1 through 5.6. The SSANOS values there show that by applying a single control limit, the in-control performances of the 20 and 1,000-CUSUM combinations are reasonably close to that of the GLR chart, especially for the 1,000-CUSUM case. In general, both CUSUM combinations have improved overall performance compared to the individual CUSUM charts – though their performance is not the best for most of the shifts, they are fairly close to the best. But note that performance of the 20-CUSUM combination is always worse than that of the 1,000-CUSUM combination or the GLR chart at every shift size. This is due to the fact that both 1,000-CUSUM combination and the GLR chart are composed of a large number of CUSUM charts; whereas only a few CUSUM charts are used in the 20-CUSUM combination. There must be some shift sizes that could not be effectively detected by the combination of only limited number of CUSUM charts. If we increase the number of CUSUM charts in the combination, we can further improve its overall performance. Ultimately, if a sufficiently large number of CUSUM charts is used, the CUSUM combination should have very similar performance to that of the GLR chart. Therefore, we see that the performance of the 1,000-CUSUM combination closely matches that of the GLR chart, both in-control and out-of-control in all Tables 5.1 through 5.6.

The same conclusion can be drawn from the other examples when $p_0 = 0.1$ and $p_0 = 0.001$. Results in Tables 5.3-5.6 show that compared to the individual CUSUM charts, the CUSUM chart combinations are better in terms of the overall performance for multiple shift sizes. If enough CUSUM charts are used in the combination, the performance will be very similar to the GLR chart.

In conclusion, performance comparisons made among the binomial GLR chart, the Shewhart chart, individual binomial CUSUM charts, the Shewhart-CUSUM combination, and combinations of multiple binomial CUSUM charts showed that the binomial GLR chart is the most effective among all the charts in detecting a wide range of shifts in the process proportion. In addition, the combined binomial CUSUM chart becomes quite similar to the binomial GLR chart if a large number of individual CUSUM charts are used for the combination, making it another candidate for having good overall performance over multiple shift sizes.

Chapter 6

Derivation and Plots of the Bernoulli GLR Chart

6.1 Derivation of the Bernoulli GLR Statistic

In the case of continuous inspection, there probably would be no natural divisions of the data stream into groups of a size $n > 1$. By artificially designating n items as a sample we might lose some useful process information and delay the time to signal if a large shift occurs (see Reynolds and Stoumbos (1999)). Therefore, it is more reasonable to construct a Bernoulli GLR chart directly based on the Bernoulli observations. The same procedure used to derive the binomial GLR statistic is used here to develop the Bernoulli GLR statistic – the only difference is that now the sample size is $n = 1$.

Recall that the maximum likelihood estimator for p_1 in the binomial case is given by

$$\hat{p}_1 = \frac{\sum_{i=\tau+1}^{k} T_i}{(k-\tau)n}.$$

When a process shifts out-of-control, if samples after the estimated change-point τ consist of items that are all defective, then $\sum_{i=\tau+1}^{k} T_i$ would be equal to $(k-\tau)n$, thus there is a chance for $\hat{p}_1$to be 1 for any value of n and $k-\tau$.

For the Bernoulli case, since the sample size is $n = 1$, T_i will either be zero if an item is conforming, or one if an item is nonconforming. Now $\hat{p}_1$ becomes

$$\hat{p}_1 = \frac{\sum_{i=\tau+1}^{k} T_i}{k-\tau}. \tag{6.1}$$

The issue with $\hat{p}_1$ is that for $n = 1$, it is very likely we will eventually have $\hat{p}_1 = 1$ for some value of $k-\tau$. If this were the case, then the Bernoulli GLR statistic

$$\lambda_{\text{Bern},k} = \ln \frac{\max\limits_{0\le\tau<k, p_0\le p_1<1} L(\tau, p_1 | T_1, T_2, \ldots, T_k)}{L(\infty, p_0 | T_1, T_2, \ldots, T_k)} = \max_{0\le\tau<k} (k-\tau)\left[\hat{p}_1 \ln \frac{\hat{p}_1(1-p_0)}{p_0(1-\hat{p}_1)} + \ln \frac{1-\hat{p}_1}{1-p_0}\right] \quad (6.2)$$

will have a non-negligible probability of being undefined, and we cannot build the Bernoulli GLR chart based on $\lambda_{\text{Bern},k}$.

To avoid having an undefined $\lambda_{\text{Bern},k}$, one way is to put some constraint such as an upper bound on $\hat{p}_1$ so that $\hat{p}_1 = 1$ would not occur. Alternatively, we could restrict the location of the estimated change-point $\hat{\tau}$ so that it cannot get too close to k – with a larger number of items used to estimate p_1, the chance that $\hat{p}_1 = 1$ would be greatly reduced. Both approaches were studied and based on simulation results shown in the next section, we found that these two restrictions had similar impacts on the performance of the corresponding Bernoulli GLR chart. In addition, it seems that by having restrictions on $\hat{\tau}$, the distribution of the corresponding GLR statistic becomes more discrete and thus more frequently the desired in-control ANOS value is not attainable. Therefore, in this dissertation we will focus on the method that build the Bernoulli GLR chart based on $\hat{p}_1'$, which is a modified MLE of p_1 with an upper bound imposed:

$$\hat{p}_1' = \min(\hat{p}_1, p_{\text{UB}}),$$

where p_{UB} is the upper bound. The Bernoulli GLR statistic based on $\hat{p}_1'$ with a moving window of size m is then given by

$$\lambda_{\text{Bern},m,k} = \begin{cases} \max\limits_{0\le\tau<k} (k-\tau)\left[\hat{p}_1' \ln \frac{\hat{p}_1'(1-p_0)}{p_0(1-\hat{p}_1')} + \ln \frac{1-\hat{p}_1'}{1-p_0}\right], & k = 1, 2, \ldots, m \\ \max\limits_{k-m\le\tau<k} (k-\tau)\left[\hat{p}_1' \ln \frac{\hat{p}_1'(1-p_0)}{p_0(1-\hat{p}_1')} + \ln \frac{1-\hat{p}_1'}{1-p_0}\right], & k = m+1, m+2, \ldots \end{cases} . \quad (6.3)$$

There are several ways to impose an upper bound on $\hat{p}_1$, among which we considered three options – a constant upper bound p_{UB}, an upper bound that is a linear function of $k-\tau$, and an upper bound whose logit function is a linear function of $k-\tau$. Simulation results (not shown here) showed that none of them was uniformly better than the others. Since a constant p_{UB} gives similar performance as the other non-constant upper bounds do but is easier to use and interpret in practice, we recommend that the Bernoulli GLR chart be constructed based on $\hat{p}_1'$ that has a constant upper bound p_{UB}.

6.2 Investigation of Two Restrictions on the Bernoulli GLR Statistic

As is explained in Section 6.1, we considered two approaches that impose different restrictions on $\lambda_{\text{Bern},k}$ so that $\lambda_{\text{Bern},k}$ would no longer be undefined, or have a very low (negligible) chance of being undefined. One way is to put an upper bound p_{UB} on $\hat{p}_1$ to ensure that the new estimate $\hat{p}_1'$ of p_1 would never be 1. The other method is to restrict the location of $\hat{\tau}$ so that it cannot get too close to k. For example, suppose the new change-point estimator, say $\hat{\tau}'$, is restricted to $\hat{\tau}' \leq k - c$, where c is a positive integer. Now that more items are always used to estimate p_1, the chance that $\hat{p}_1 = 1$ would be greatly reduced. Both approaches were investigated and the corresponding simulation results are presented below in Table 6.1.

In Table 6.1, we considered the case when $p_0 = 0.01$ and selected the control limits so that the in-control ANOS values of all the charts are close to 29,350, which is equal to the in-control ANOS value for the binomial GLR chart studied previously in Chapter 5 with $p_0 = 0.01$ and $n = 100$. The window size m_B is set to be 30,000 for all the charts, and the change-point τ is 10,000. Columns [2]-[4] are SSANOS values for three GLR charts with different values of p_{UB} so that the estimate of p_1 is restricted to be no greater than p_{UB}, whereas columns [5]-[10] are SSANOS values based on GLR charts with values of c ranging from 15 to 90, and the restriction is on the estimate of τ so that it is no larger than $k - c$. For any value of c that is smaller than 15, we found it impossible to adjust the control limit of the GLR chart to achieve an in-control ANOS close to 29,350.

TABLE 6.1. SSANOS values for the two Bernoulli GLR chart with different restrictions when $p_0 = 0.01$.

	Bernoulli GLR with p_{UB}			Bernoulli GLR with $\hat{\tau}' \leq k$-c					
	p_{UB}=0.025	p_{UB}=0.04	p_{UB}=0.1	c=15	c=25	c=35	c=50	c=70	c=90
p	h=4.94	h=5.85	h=7.01	h=7.13	h=7.13	h=6.50	h=6.50	h=6.04	h=6.04
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.010	29241.53	29136.98	28947.65	29234.39	29195.54	29248.05	29259.88	29331.16	29393.95
0.015	2377.17	2720.52	3262.92	3431.78	3429.46	3105.45	3105.86	2892.87	2890.24
0.020	863.30	951.96	1118.79	1170.75	1170.10	1068.60	1067.66	1003.51	1003.63
0.025	486.57	514.99	592.71	615.87	615.64	567.60	567.21	537.91	538.35
0.030	331.53	338.08	380.18	392.92	392.89	365.08	364.57	350.01	349.57
0.040	200.35	192.26	206.73	211.69	211.97	200.10	200.18	196.43	196.60
0.050	143.06	132.02	135.92	138.70	138.56	133.41	133.48	134.54	134.48
0.060	111.26	100.06	99.50	100.94	100.95	98.94	98.93	102.34	102.20
0.070	91.09	80.41	77.45	78.29	78.32	78.42	78.33	83.08	83.07
0.080	76.97	67.27	63.07	63.63	63.65	64.87	64.98	70.21	70.13
0.090	66.70	57.75	52.97	53.29	53.32	55.43	55.48	61.02	61.00
0.100	58.78	50.55	45.58	45.76	45.77	48.54	48.47	54.07	54.03
0.150	36.68	31.15	26.71	26.73	26.71	30.33	30.35	34.74	34.76
0.200	26.63	22.61	18.98	18.95	18.96	22.25	22.26	25.65	25.64
0.300	17.25	14.64	12.11	12.12	12.12	14.47	14.48	16.73	16.71
0.500	10.03	8.49	6.98	6.99	6.99	8.40	8.40	9.74	9.74
0.750	6.48	5.46	4.46	4.47	4.47	5.40	5.40	6.30	6.29
1.000	4.72	3.96	3.21	3.22	3.22	3.92	3.92	4.58	4.58

We see from columns [5] and [6] that when c equals to 15 or 25, the two GLR charts have very similar performances both in-control and out-of-control, with their control limits both being 7.13. In fact more simulations were done for values of c that are between 15 and 25 (results not shown here), and we found that as long as c is between 15 and 25, different values of c would have the same effect on the performance of the GLR chart, and the control limit remains the same as well. However, if $c < 15$ or $25 < c < 35$, an in-control ANOS close to 29,350 is no longer attainable, and the corresponding SSANOS values would also significantly change. This same pattern can be found for more ranges of c such as those between 35 and 50, or between 70 and 90. It seems that the distribution of the GLR statistic becomes more discrete when a restriction is imposed on $\hat{\tau}$, which makes it harder to find the desired in-control ANOS value.

Moreover, across columns [5] through [10], we find that as c increases, the GLR charts tend to be more effective in detecting small shifts, but less effective for large shifts. There is no "optimal" value of c that we could use to design a GLR chart that performs well for all shift sizes.

Meanwhile, the results in columns [2] through [4] for SSANOS values of the GLR chart based on different values of p_{UB} show a similar pattern as what we have seen from columns [5] through [10]: a large p_{UB} requires a large control limit, which makes it harder to detect small shifts; however, the large p_{UB} allows for faster detection of large shifts. There is no value of p_{UB} that leads to a GLR chart with uniformly better overall performance.

In addition, notice that if we compare column [2] to [9], [3] to [7], and [4] to [5], the two GLR charts with different restrictions have almost the same SSANOS values for detecting large shifts; however, the GLR chart with the upper bound has quicker detection of small shifts compared to the GLR chart with $\hat{\tau}' \leq k - c$. Therefore, we decided to use the upper bound approach based on the modified Bernoulli GLR statistic so that $\lambda_{\mathrm{Bern},k}$ would no longer be undefined, which is given in Equation (6.3) in the previous section.

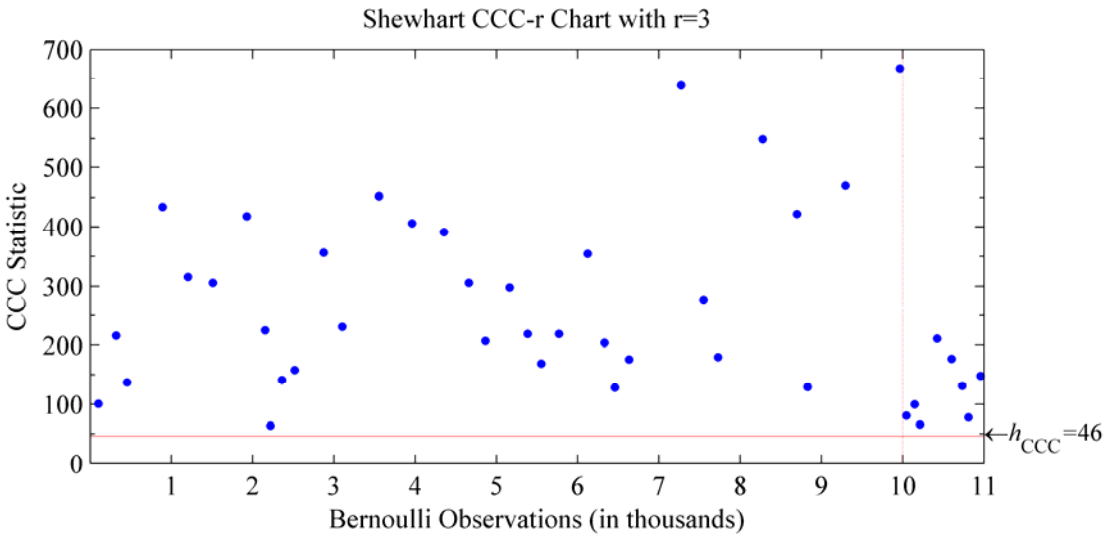
6.3 Example Plots of the Bernoulli GLR Chart

As an example, we first simulated 10,000 observations from the Bernoulli distribution with $p_0 = 0.01$ to show what the data look like when the process is in-control. After the first 10,000 observations, 1,000 more Bernoulli observations were generated from the same process but with p increased from 0.01 to 0.025, so that now the process is regarded as out-of-control. The true process change-point is $\tau = 10{,}000$. In particular, two values are considered for p_{UB} in the GLR chart and p_1 in the CUSUM chart. One value is for detecting a small shift with $p_{\mathrm{UB}} = p_1 = 0.02$; another is for detecting a large shift with $p_{\mathrm{UB}} = p_1 = 0.1$.

In Figure 6.1, a Shewhart CCC-r chart is plotted with $r = 3$ and $h_{\mathrm{CCC}-3} = 46$. This CCC-3 chart is constructed based on the cumulative count of conforming items until a third nonconforming item is observed. Fewer counts of cumulative conforming items indicate process quality deterioration, thus an increase in p. Therefore, only the lower control limit is plotted and any point falling below it triggers a signal. Notice that there are no signals (false-alarms) in the first 10,000 in-control observations. After the increase in p at $\tau = 10{,}000$ we can see that the cumulative counts of conforming items seem to be smaller, but there is no point below the control limit in the next 1,000 observations. Since the shift size considered here is small, it is not

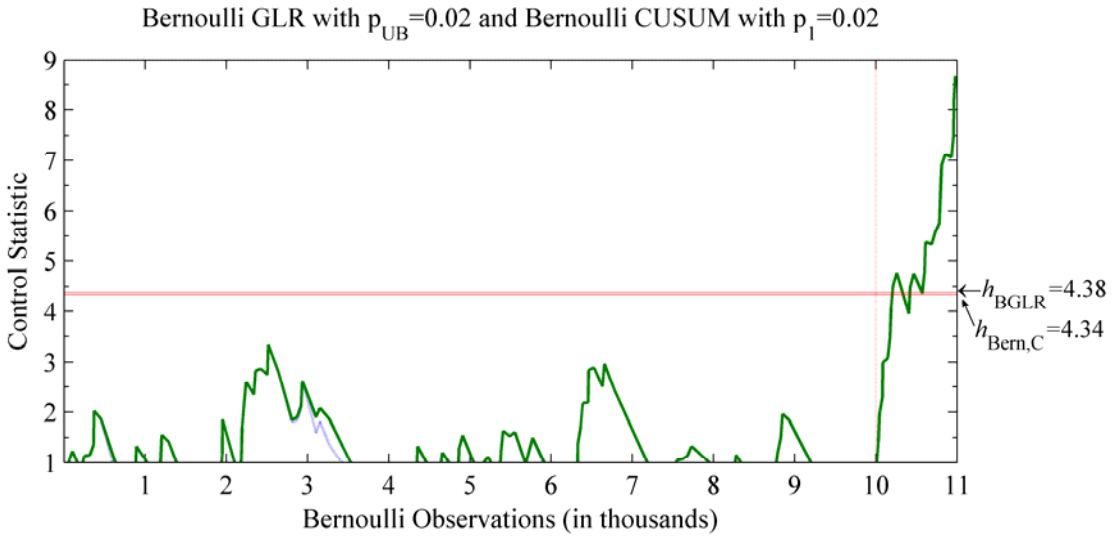
surprising that the Shewhart CCC-3 chart is not able to detect this shift within 1,000 out-of-control observations.

FIGURE 6.1. The Shewhart CCC-3 chart for 10,000 in-control observations with $p_0 = 0.01$ and 1,000 out-of-control observations with $p = 0.025$. No signal is given.



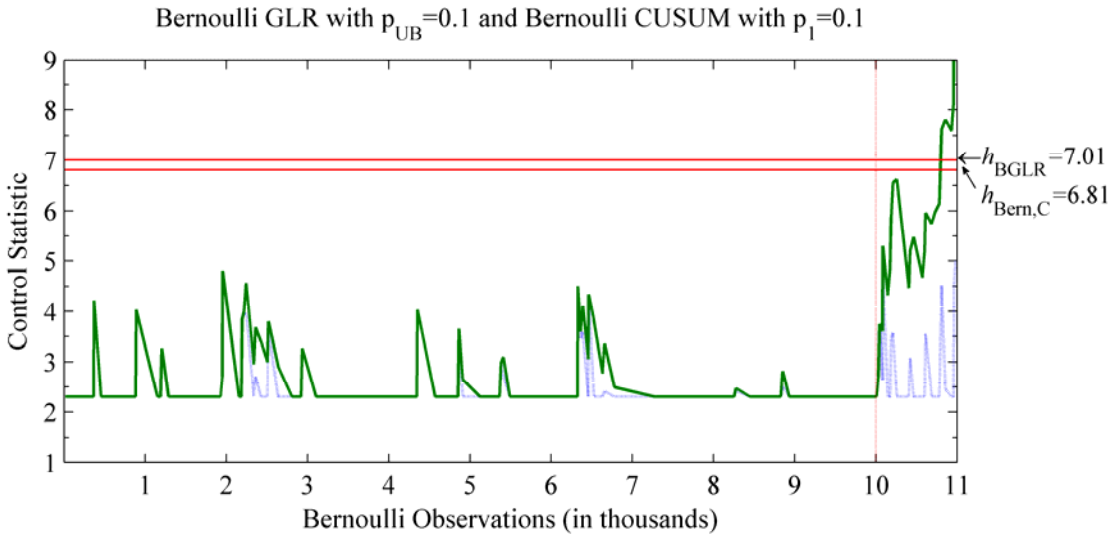
In Figure 6.2, the Bernoulli GLR chart and Bernoulli CUSUM chart are plotted on the same plot with $p_{\mathrm{UB}} = p_1 = 0.02$. The parameter p_{UB} of the GLR chart has the same value as the parameter p_1 of the CUSUM chart so that both charts would have similar out-of-control performance. The solid line denotes the GLR chart, with control limit $h_{\mathrm{BGLR}} = 4.38$ and $m_{\mathrm{B}} = 30{,}000$, and the dashed line is for the CUSUM chart, with control limit $h_{\mathrm{BernC}} = 4.34$. Notice that since we only simulated 11,000 observations in total here, there was actually no window used for the GLR chart in this plot. Neither of the two charts gives a false-alarm within the first 10,000 in-control observations. The pattern of the CUSUM chart looks very similar to the pattern of the GLR chart, and both of them signal at observation 10,212, giving the same value of their control statistics of 4.515. In addition, at observation 10,212, the estimate of the change-point provided by the GLR chart is $\hat{\tau} = 9{,}964$, which is very close to the actual change-point $\tau = 10{,}000$. The estimated shift size $\hat{p}_1$ is 0.02 due to the restriction imposed by $p_{\mathrm{UB}} = 0.02$. In this example, although the true shift size is not exactly 0.02, both the Bernoulli GLR and CUSUM charts are able to detect such a small shift within relatively few out-of-control observations.

FIGURE 6.2. Plots of Bernoulli GLR and CUSUM charts for 10,000 in-control observations with $p_0 = 0.01$ and 1,000 out-of-control observations with $p = 0.025$, where $p_{\text{UB}} = p_1 = 0.02$. The out-of-control signals are given at observation 10,212.



Another value of $p_{\text{UB}} = p_1 = 0.1$ is considered in Figure 6.3 for the Bernoulli GLR and CUSUM charts based on the same data shown in Figures 6.1 and 6.2. Again, the solid line denotes the GLR chart, with control limit $h_{\text{BGLR}} = 7.01$ and $m_{\text{B}} = 30{,}000$, and the dashed line is for the CUSUM chart, with control limit $h_{\text{BernC}} = 6.81$. No false-alarm is given when the process is in-control. This time, though the in-control patterns are still similar for the GLR and the CUSUM charts, the out-of-control behavior of the CUSUM chart is quite different from the GLR chart. Since the value of $p_1 = 0.1$ is far from the true shift size 0.025, the CUSUM chart is not able to pick up this shift within 1,000 out-of-control observations. On the other hand, the GLR chart with $p_{\text{UB}} = 0.1$ is able to detect this small shift of 0.025, although a longer time is needed for it to signal – it takes the GLR chart 809 observations to signal here but only 212 observations to signal in the previous situation. At observation 10,809, the GLR statistic is 7.612, and the estimates of τ and p are $\hat{\tau} = 9{,}964$ and $\hat{p}_1 = 0.026$, respectively – both of them closely match the corresponding true values.

FIGURE 6.3. Plots of Bernoulli GLR and CUSUM charts for 10,000 in-control observations with $p_0 = 0.01$ and 1,000 out-of-control observations with $p = 0.025$, where $p_{\text{UB}} = p_1 = 0.1$. Only the GLR chart signals at observation 10,809.



We find from the above plots that in general the Shewhart-type chart is not effective in detecting small shifts. If the true shift size is in the neighborhood of the values of p_{UB} and p_1 in the GLR and CUSUM charts, then both of them have satisfactory performance in detecting that shift. Figure 6.3 shows that when the true shift size is much smaller than what is specified for p_{UB} or p_1, the performance of the CUSUM chart would be significantly affected and thus no longer effective for detect that shift, while the GLR chart is still able to detect that shift but with a longer signal time. In fact, with the CUUSM chart it seems reasonable to assume that we are expecting a shift close to p_1, but the idea behind the GLR chart is that we are not expecting any particular shift size. Here, if a particular upper bound p_{UB} is used in the GLR chart, then we are expecting a shift within $(p_0, p_{\text{UB}}]$. Furthermore, as we would discuss later in Chapter 8, that the Bernoulli CUSUM chart is actually a special case of the Bernoulli GLR chart with the upper bound, so it is not surprising to see that the control statistic of the GLR chart is always at least as large as the control statistic of the corresponding CUSUM chart, and h_{BGLR} is always slightly larger than h_{BernC} as well.

Chapter 7

Designing the Bernoulli GLR Chart

Unlike the binomial GLR chart where only two parameters (the window size m and the control limit h_{GLR}) need to be specified for any practical application, the Bernoulli GLR chart introduces a third parameter, the upper bound p_{UB} that we put on $\hat{p}_1$ to prevent the log likelihood ratio statistic $\lambda_{\mathrm{Bern},m,k}$ from being undefined. Thus, to design a Bernoulli GLR chart, we need to specify three parameters – m_{B}, h_{BGLR}, and p_{UB}.

7.1 Selecting the Window Size

Simulation results (not shown here) confirmed that the window size m_{B} in a Bernoulli GLR chart can be determined in the same way as m for a binomial GLR chart, i.e. a moving window of size m_{B} that equals the in-control ANOS is sufficiently large to obtain almost the same performance as the GLR chart without any window, in terms of detecting even a very small shift. This rule of thumb will be followed here and we will pick an m_{B} value that is close to the desired in-control ANOS value.

7.2 Selecting the Upper Bound

Various values of p_{UB} ranging from small to large were investigated and we found that the role of p_{UB} for the GLR was actually similar to the role of the turning parameter p_1 in the CUSUM chart. Choosing a proper value for p_{UB} depends on the specific shift size one wants to detect. For example, if a small shift is anticipated, we should use a small value of p_{UB}; for detecting a moderate or large shift, it is necessary to use a relatively large value for p_{UB}. In general, the GLR chart with a particular p_{UB} has good performance in detecting any shift that falls in $(p_0, p_{\mathrm{UB}}]$.

7.3 Selecting the Control Limit

For continuous inspections without any grouping, the sample size n is always equal to 1. Selection of the control limit h_{BGLR} now depends on p_0, the desired in-control ANOS (ICANOS) value, as well as the upper bound p_{UB}. Moreover, based on simulation results, we found that h_{BGLR} can be determined as long as the product of ICANOS and p_0, and the ratio of p_{UB} and p_0 are known. That is to say, h_{BGLR} only depends on the expected number of nonconforming items in the time period before a false-alarm when the process is in-control (ICANOS$\times p_0$), and the size of p_{UB} relative to p_0 (p_{UB}/p_0).

Table 7.1 tabulates h_{BGLR} when the desired in-control expected number of nonconforming items ranges from 100 to 1,000, based on five ratios of p_{UB} and p_0 that are equal to 2, 3, 5, 10 and 20. As was discussed in Section 7.2, for a specific Bernoulli GLR chart, the value of p_{UB} depends on the shift size of interest. For instance, suppose the objective is to detect a shift from $p_0 = 0.05$ to $p = 0.15$, then we should set p_{UB} to be 0.15, which corresponds to $p_{\mathrm{UB}}/p_0 = 3$. In general, for detecting small shifts, a value of 2 or 3 is often considered for p_{UB}/p_0; $p_{\mathrm{UB}}/p_0 = 5$ can be used when we want to detect moderate shifts; and a large value of p_{UB}/p_0 such as 10 or 20 should be used if the goal is to detect very large shifts. In this sense, Table 7.1 can be used to help design a Bernoulli GLR chart that works for many practical applications.

As an example, consider a situation that $p_0 = 0.05$ and a user want to design a Bernoulli GLR chart for detecting a relatively small shift size that is roughly three times as large as p_0. If an ICANOS value of 10,000 is desired, it can be easily found from Table 7.1 that an h_{BGLR} value that corresponds to ICANOS$\times p_0 = 500$ and $p_{\mathrm{UB}}/p_0 = 3$ should be selected, which is equal to 5.90. Through simulation we found an $h_{\mathrm{BGLR}} = 5.90$ gives an ICANOS value of 9958.78, which is very close to the desired value of 10,000.

TABLE 7.1. h_{BGLR} for Bernoulli GLR chart with different combinations of ICANOS$\times p_0$ and p_{UB}/p_0.

ICANOS x p_0	p_{UB} / p_0 =				
	2	3	5	10	20
	$p_0\in$[0.001, 0.1]	$p_0\in$[0.001, 0.1]	$p_0\in$[0.001, 0.1]	$p_0\in$[0.001, 0.05]	$p_0\in$[0.001, 0.02]
100	3.37	4.26	5.07	5.97	6.45
150	3.75	4.67	5.50	6.35	7.04
200	4.02	4.96	5.81	6.62	7.41
250	4.24	5.18	6.02	6.82	7.65
300	4.42	5.37	6.20	7.03	7.88
350	4.57	5.53	6.36	7.22	8.00
400	4.71	5.67	6.52	7.37	8.12
450	4.83	5.79	6.64	7.50	8.23
500	4.93	5.90	6.76	7.62	8.35
550	5.03	6.00	6.85	7.73	8.41
600	5.12	6.09	6.94	7.81	8.49
650	5.20	6.17	7.02	7.90	8.57
700	5.27	6.25	7.11	7.98	8.61
750	5.34	6.32	7.18	8.06	8.68
800	5.41	6.39	7.26	8.13	8.74
850	5.47	6.45	7.32	8.18	8.79
900	5.53	6.51	7.36	8.25	8.85
950	5.59	6.57	7.43	8.28	8.89
1000	5.63	6.63	7.50	8.35	8.98

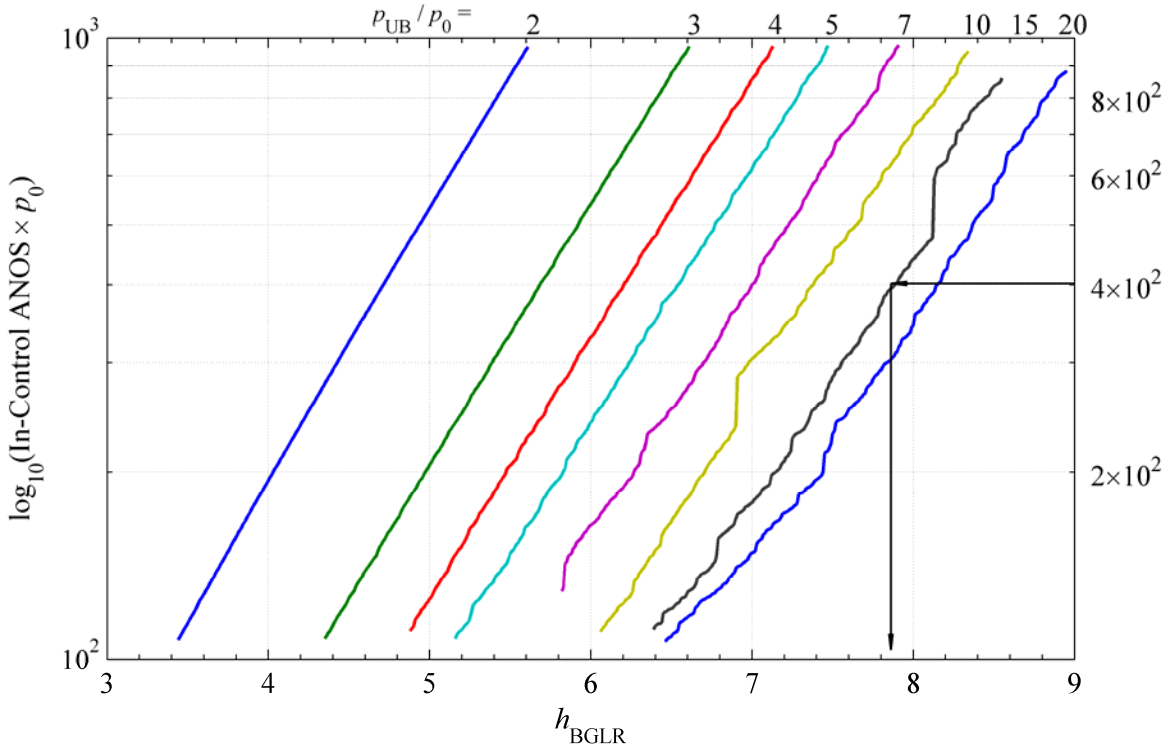
Notice that although theoretically the value of an upper bound p_{UB} could be as large as 0.99, we would recommend not using such a large value due to the discreteness of the Bernoulli distribution. All the upper bounds used in Table 7.1 are set to be less than or equal to 0.5. If the in-control proportion p_0 is small, then using p_{UB} as large as 0.5 is sufficient to detect a quite large shift. If p_0 itself is large, say, $p_0 > 0.1$, then the relationship between ICANOS$\times p_0$ and h_{BGLR} becomes less linear and large jumps would be expected in the ICANOS values because the issue of discreteness of the Bernoulli distribution turns out to be more severe for large values of p_0. In this case, simulation should be used to find h_{BGLR}. In fact, in Table 7.1, we only consider $0.001 \le p_0 \le 0.1$ for $p_{\mathrm{UB}}/p_0 = 2$, 3 and 5, $0.001 \le p_0 \le 0.05$ for $p_{\mathrm{UB}}/p_0 = 10$, $0.001 \le p_0 \le 0.02$ for $p_{\mathrm{UB}}/p_0 = 20$, and the values of ICANOS$\times p_0$ ranging from 100 to 1,000. Later in this section, one can find more combinations of ICANOS$\times p_0$ and p_{UB}/p_0 that are

presented in Figure 7.1. However, for any value of p_0, p_{UB}/p_0, or ICANOS$\times p_0$ that is not tabulated in Table 7.1 or plotted in Figure 7.1, users of the Bernoulli GLR chart should perform simulation to find the appropriate control limit.

In general, for a given p_{UB}, we found that h_{BGLR} is approximately linearly related to log10 of the product of ICANOS and p_0. This linear relationship is illustrated in Figure 7.1 for various cases with the ratio of p_{UB} and p_0 ranging from 2 to 20. Users can also use this chart to find h_{BGLR} for some combinations of ICANOS$\times p_0$ and p_{UB}/p_0 that are not given in Table 7.1, such as $p_{\mathrm{UB}}/p_0 = 4$, 7, or 15. As an example, consider monitoring a Bernoulli process with $p_0 = 0.005$. Suppose an ICANOS value of 80,000 is desired, and our interest is to effectively detect a relatively large shift, say, $p_{\mathrm{UB}}/p_0 = 15$. From Figure 7.1 we first find the intersect of a horizontal line that labeled 4×10^2 $(= 80{,}000 \times 0.005)$ and the line labeled $p_{\mathrm{UB}}/p_0 = 15$, then from that point if we draw a vertical line down to the h_{BGLR} axis, the new intersect would give the value of h_{BGLR} to be used in this particular chart, which is roughly equal to 7.82. Through simulation we found an $h_{\mathrm{BGLR}} = 7.82$ gives an ICANOS value of 83215.52, which is reasonably close to the desired value of 80,000.

Again, all the lines plotted in Figure 7.1 are based on p_{UB} values no larger than 0.5. Also, depending on the ratio of p_{UB} and p_0, the value of p_0 should be always restricted to the range of 0.001 and 0.1. If these conditions are satisfied, the control limit h_{BGLR} of a Bernoulli GLR chart can be easily obtained from Figure 7.1, for any desired in-control expected number of nonconforming items before a false-alarm between 100 and 1,000.

FIGURE 7.1. Plot of h_{BGLR} and $\log_{10}$(ICANOS$\times p_0$) for eight values of p_{UB}/p_0 that range from 2 to 20, with $p_0 \in [0.001, 0.1]$ if $p_{\mathrm{UB}}/p_0 = 2, 3, 4, 5, 7$; $p_0 \in [0.001, 0.05]$ if $p_{\mathrm{UB}}/p_0 = 10$; and $p_0 \in [0.001, 0.02]$ if $p_{\mathrm{UB}}/p_0 = 15$ or 20.



Chapter 8

Bernoulli Control Chart Performance Comparisons

8.1 Comparing the Bernoulli GLR Chart with Other Charts

In this chapter, we compare the overall performance of various control charts based on the continuous inspection, which, as was defined in Section 2.1, assumes that a continuous stream of inspected items is available all the time. Particularly, in this section, we evaluate performance of the Bernoulli GLR chart relative to the Shewhart-type CCC-*r* chart, the Bernoulli CUSUM chart, the Bernoulli EWMA chart and the binomial GLR chart, in detecting increases in p. As is pointed out by Szarka and Woodall (2011) in their review paper of control charts for monitoring Bernoulli processes, "… Due to well-established theory, it is very difficult to compete with the overall steady-state statistical performance of the Bernoulli CUSUM chart… We strongly recommend that the Bernoulli CUSUM chart be included in all comparisons", here we will mainly focus on the comparisons between the Bernoulli GLR chart and the Bernoulli CUSUM chart. For the purpose of fair comparisons, all the charts are set up to have approximately the same in-control performance. Comparisons are made across a wide range of shift sizes for p when three levels of p_0 are considered.

Previous studies have shown that the Shewhart-type charts may detect large shifts very quickly, but are insensitive to small shifts; the Bernoulli CUSUM chart is good around its tuned shift size p_1, but performs worse for other shifts away from p_1; performance of the Bernoulli EWMA chart depends on the value of its smoothing parameter λ. On the other hand, the binomial GLR chart, which is based on artificially aggregating items into samples of size $n > 1$, uses the maximum likelihood method to estimate the shift size directly from the data so that its performance is satisfactory for a wide range of shift sizes and does not depend on p_1. However, by grouping items into samples, it becomes less effective to signal a large shift in a timely manner. Moreover, for the binomial GLR chart with grouping, any estimate of τ will be between

samples while it is likely that τ occurs within a sample. This may not be much of a problem if n is small, but could add to the bias and mean squared error of $\hat{\tau}$ for large values of n. Alternatively, with an upper bound p_{UB} imposed on $\hat{p}_1$ to prevent it from being 1, an extra charting parameter is introduced to the Bernoulli GLR chart, which makes it behave very much like the Bernoulli CUSUM chart – it has good performance when the actual shift size p satisfies $p_0 < p \leq p_{\mathrm{UB}}$, but is not very effective if $p > p_{\mathrm{UB}}$.

For the first comparison, consider the case of $p_0 = 0.01$. Suppose a change occurs in a Bernoulli process after item 10,000 is observed, i.e. $\tau = 10{,}000$. A window of size $m_{\mathrm{B}} = 30{,}000$ is used to ensure that the Bernoulli GLR chart performs similarly as if no window is used. In order to compare the out-of-control performance of the Bernoulli-type charts to the binomial-type charts, we select all the control limits h_{BGLR}, h_{BernC} and h'_{BernE} such that the in-control ANOS values of the corresponding Bernoulli GLR, CUSUM and EWMA charts are close to 29,350, which is the in-control ANOS value of the binomial GLR chart with $n = 100$ and $h_{\mathrm{GLR}} = 4.13$. The values of $h_{\mathrm{CCC}-2}$, $h_{\mathrm{CCC}-3}$, and $h_{\mathrm{CCC}-4}$ for three Shewhart CCC-r charts are selected in the same manner to match the in-control ANOS value of 29,350. Since $h_{\mathrm{CCC}-r}$ can only be integer values, the actual in-control ANOS value of the CCC-r chart might be somewhat different from the desired value.

In Table 8.1, columns [2]-[4] give SSANOS values of the Shewhart CCC-r charts, with three different values of r ($r = 2, 3,$ and 4) being used. Columns [5]-[8] give SSANOS values of four Bernoulli GLR charts with p_{UB} being 0.02, 0.05, 0.1, and 0.2, respectively. We choose these four values of p_{UB} to cover different sizes of shift, with each Bernoulli GLR chart being able to detect a particular shift range. Columns [9]-[12] are the SSANOS values of four Bernoulli CUSUM charts with $p_1 = 0.02$, 0.05, 0.1, and 0.2, respectively. The four values of p_1 are selected to be the same as those of p_{UB}, so that we can fairly compare the performance of each pair of the GLR and CUSUM charts with $p_{\mathrm{UB}} = p_1$.

Columns [13]-[15] of Table 8.1 are SSANOS values for the Bernoulli EWMA chart with the smoothing parameter λ being 0.00001, 0.0001, and 0.001, respectively. In general, a smaller value of λ is used to detect small shifts, while a large value of λ is selected for detecting large shifts. As we have discussed in Section 2.4.3, the EWMA chart presented here is based on the

approximate control limit and with a built-in reset scheme. The last two columns [16] and [17] contain SSANOS values of the two binomial GLR charts discussed before with sample size $n = 16$ and $n = 100$, respectively. However, instead of assuming that the shift can only occur between samples (as we did when we were investigating the binomial GLR charts in Chapters 3-5), here it is more reasonable to allow the shift to occur within a sample consisting of observations from continuous inspection that have been artificially grouped into samples of $n > 1$.

We see from Table 8.1 that for the Shewhart CCC-r charts, when a small value of r such as $r = 2$ is used, it has good performance in detecting large shifts, but is not effective when the shift size is small or moderate. To improve its performance for detecting smaller shifts, one could select a larger value of r, such as $r = 4$. However, since now more nonconforming items are required to trigger a signal, it delays the signal time if a large shift occurs. In general, when compared to other charts in Table 8.1, we find that the overall performance of this Shewhart-type chart is not very competitive relative to the GLR, CUSUM, and EWMA charts.

Notice that in Table 8.1, the four Bernoulli GLR charts perform well, especially for small shifts. For example, if we compare column [5] to [9], which are based on the Bernoulli GLR and CUSUM charts both designed to detect small shifts, we find that for $p < 0.02$, the GLR chart with $p_{\mathrm{UB}} = 0.02$ has much better performance than the CUSUM chart with $p_1 = 0.02$; at $p = 0.02$, the GLR chart gives an SSANOS of 836.64, which is still smaller than that of the CUSUM chart (839.41). For shifts greater than 0.02, the GLR chart has only slightly larger SSANOS values than the CUSUM chart. This advantage of the Bernoulli GLR chart over the Bernoulli CUSUM chart that the former has quicker detection of small shifts also exists if other values of p_{UB} are used for the charts to effectively detect moderate or large shifts. If we compare column [6] to [10], [7] to [11], and [8] to [12], we find that for a given pair of the GLR and CUSUM charts with $p_{\mathrm{UB}} = p_1$, the signal time of the GLR chart is only slightly longer than that of the CUSUM chart at all shifts $p \geq p_1$ (or p_{UB}), but for $p < p_1$ (or p_{UB}), the GLR chart is much better than the corresponding CUSUM chart.

The three Bernoulli EWMA charts in Table 8.1 are set up based on the reset scheme. In column [13], the Bernoulli EWMA chart with $\lambda = 0.0001$ has the best performance among all

the charts for detecting small shifts such as $p < 0.015$, but it is insensitive to moderate or large shifts. If a larger value of λ is used, the EWMA chart has improved performance in terms of detecting large shifts, but is not as good for detecting small shifts. None of these three EWMA charts is uniformly better than any of the GLR or CUSUM charts. We also found that without using the reset scheme, the steady state performance of an EWMA chart would be worse (results not shown here).

Compared to the Bernoulli-type charts in Table 8.1, the two binomial GLR charts in columns [16] and [17] have reasonably good performance for detecting small shifts, especially for the chart with $n = 100$. However, since those binomial charts are based on artificially grouping items into samples of size $n > 1$, a decision about the process cannot be made until all n items from a sample are inspected, and having a large sample size such as $n = 100$ would significantly delay the signal time when a large shift occurs. The binomial GLR chart with a smaller sample size such as $n = 16$ (column [16]) has relatively satisfactory performance for detecting large shifts, but is not as effective as the GLR chart with $n = 100$ for detecting small shifts. Since now we allow the shift to occur within a sample, which is a more reasonable assumption than assuming that the shift can only occur between samples, the SSANOS values of the binomial charts approach $n/2$ when the shift size is very large.

For another example, consider the monitoring of a high quality process with $p_0 = 0.001$. Table 8.2 gives SSANOS values for the same charts listed in Table 8.1, but with modified parameters. Here the change-point τ is set to be 100,000. A moving window of size $m_{\mathrm{B}} = 300{,}000$ is used for the Bernoulli GLR chart. Again, three values of $r = 2, 3,$ and 4 are considered for the Shewhart CCC-r chart. Control limits for all of the charts are selected in the same way as in the previous example.

The results in Table 8.2 show that the Bernoulli GLR chart is very effective in detecting small shifts. For the other shift sizes, its performance is only a bit worse than the corresponding Bernoulli CUSUM chart where $p_1 = p_{\mathrm{UB}}$. The Shewhart CCC-r charts perform poorly in general. The Bernoulli EWMA chart can be tuned to have the best performance in detecting small shifts (e.g. $p \leq 1.5p_0$) by selecting an extremely small λ, but then it is not effective if a large shift occurs. The binomial GLR chart with $n = 1{,}000$ in column [17] has good performance when the

shift size is small, but is bad for large shifts since it cannot signal until the whole sample is inspected, which delays its signal time substantially. Meanwhile, the binomial GLR chart with $n = 150$ in column [16] performs reasonably well unless the shift size is quite large, say $p \geq 0.05$, which actually is not very likely to happen in practice when $p_0 = 0.001$. From this we see that using $n = 150$ instead of $n = 1{,}000$ gives better performance for large shifts, at the expense of worse performance for small shifts. Therefore, for a binomial chart to have better detection of large shifts, one could use a smaller sample size, but that presumably would make it harder to detect small shifts. In fact, no sample size smaller than 150 is considered here for the situation in which a binomial GLR chart is used to monitor a high quality process based on the continuous inspection.

Finally, consider monitoring a process with a higher in-control nonconforming rate of $p_0 = 0.1$. Table 8.3 gives SSANOS values for the same charts listed in Table 8.1 and 8.2, but with modified parameters. Here we assume the change-point $\tau = 2{,}000$. A moving window of size $m_{\mathrm{B}} = 6{,}000$ is used for the Bernoulli GLR chart. Since p_0 is relatively large in this case, only r values of 3 and 4 are considered for the Shewhart CCC-r charts. Again, we select all the control limits in the same fashion as before. The conclusion we draw from Table 8.3 is similar to what we have seen in Tables 8.1 and 8.2: the Bernoulli GLR chart is found to be very effective in detecting small shifts, and is only slightly inferior to the corresponding Bernoulli CUSUM chart with $p_1 = p_{\mathrm{UB}}$ for detecting other shift sizes. The Shewhart CCC-r chart is preferred if the primary interest is to detect large shifts; the binomial GLR chart with a relatively large sample size is better if small shifts are expected to occur. The Bernoulli EWMA chart can be tuned to be extremely good for small shifts, but cannot be effective in detecting moderate or large shifts at the same time.

In summary, for monitoring Bernoulli processes, the results presented in Tables 8.1-8.3 show that the Shewhart-type CCC-r charts should only be considered for detecting large shifts. In general we do not recommend using charts based on artificial subgrouping in the situation of continuous inspection, unless a small sample size is used and relatively quick detection of small shifts can be maintained. Whether to choose a Bernoulli EWMA chart over a Bernoulli GLR or CUSUM chart largely depends on one's personal preference – at a particular smoothing level, the EWMA cannot beat any of the GLR or CUSUM charts in terms of its overall performance,

even with a built-in reset feature. The ability of the Bernoulli GLR chart with the upper bound to detect a wide range of shifts is slightly better than that of the Bernoulli CUSUM chart. For these two charts with $p_{\mathrm{UB}} = p_1$, although they both have fast detection in the neighborhood of the target shift sizes that they were designed for, the Bernoulli GLR chart is much better than the Bernoulli CUSUM chart in detecting small shifts with $p < p_{\mathrm{UB}}$ (or p_1). When p is much larger than $p_{\mathrm{UB}} = p_1$, both charts do not perform well, even though the CUSUM chart is slightly better than the GLR chart.

In terms of the effort required for designing the various chart mentioned above, the Shewhart CCC-r chart is easier to design with only two parameters r and $h_{\mathrm{CCC}-r}$ that need to be specified. The Bernoulli GLR, CUSUM, and EWMA charts all have three charting parameters. However, for the Bernoulli GLR chart, it is fairly straightforward to select the window size m_{B}, and the upper bound p_{UB} can be chosen the same way as p_1 in the CUSUM chart. Moreover, the Bernoulli GLR control limit h_{BGLR} can be easily obtained from Table 7.1 or Figure 7.1 for any $0.001 \leq p_0 \leq 0.1$, $p_{\mathrm{UB}} \leq 0.5$, and 100 < ICANOS$\times p_0$< 1,000. On the other hand, selecting parameters for the Bernoulli CUSUM or EWMA charts is never a trivial job and requires more effort than designing a Bernoulli GLR chart. Thus, for monitoring Bernoulli processes, we would recommend applying the Bernoulli GLR chart since it is not only simpler to design such a chart for practical applications, but also this chart has better performance over most of its competitors when the focus is on effectively detecting small shifts.

TABLE 8.1. SSANOS values for various control charts with $p_0 = 0.01$.

	Shewhart CCC-r			Bernoulli GLR				Bernoulli CUSUM				Bernoulli EWMA			Binomial GLR	
	r=2	r=3	r=4	p_{UB}=0.02	p_{UB}=0.05	p_{UB}=0.1	p_{UB}=0.2	p_1=0.02	p_1=0.05	p_1=0.1	p_1=0.2	λ=0.0001	λ=0.001	λ=0.01	n=16	n=100
p	h=14	h=46	h=93	h=4.38	h=6.19	h=7.01	h=7.84	h=4.34	h=6.04	h=6.81	h=7.51	h=0.0113	h=0.0169	h=0.0453	h=6.29	h=4.13
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]
0.010	27708.00	29049.46	28879.67	29204.97	29235.49	28947.65	29252.84	29343.92	29553.50	28889.38	29220.21	29254.62	29460.54	29357.40	31410.39	29350.00
0.013	12948.30	11222.34	9640.38	4457.73	5851.55	6672.54	7668.67	4918.60	8489.10	10579.94	12772.57	3548.85	4586.29	8965.25	6309.77	4682.18
0.015	8577.28	6772.89	5473.00	2222.50	2871.64	3260.42	3745.27	2360.58	4441.76	6184.14	8139.87	2035.11	2231.47	4855.73	3070.94	2358.46
0.020	3777.64	2549.82	1868.27	836.64	995.56	1117.78	1269.77	839.41	1380.21	2178.54	3299.77	968.37	826.85	1559.18	1066.28	902.17
0.025	2028.93	1244.31	888.88	489.97	534.43	593.13	667.17	487.89	645.24	1020.21	1656.74	633.90	488.97	726.50	576.19	533.90
0.030	1231.15	717.34	518.06	343.62	348.16	380.46	424.79	341.25	383.38	576.96	956.63	471.02	345.28	424.85	376.28	378.99
0.040	573.40	324.42	250.52	214.55	194.45	207.21	228.31	212.88	198.02	259.65	417.75	311.26	216.80	210.34	211.46	246.40
0.050	324.50	189.13	159.11	155.80	131.63	136.49	148.94	154.82	131.00	153.55	230.87	232.33	158.14	134.89	144.70	188.57
0.060	207.87	127.69	116.46	122.47	98.88	100.00	107.78	121.58	97.45	105.92	147.70	185.57	124.46	98.20	109.09	156.73
0.070	145.09	94.86	92.70	100.91	78.91	77.91	83.12	100.14	77.72	79.96	104.09	154.40	102.59	77.04	87.37	136.37
0.080	107.32	75.16	77.71	85.84	65.68	63.58	67.08	85.10	64.61	63.95	78.61	132.30	87.22	63.33	73.18	122.58
0.090	83.56	62.27	67.36	74.63	56.22	53.46	55.88	73.99	55.27	53.22	62.57	115.68	75.96	53.70	62.74	112.85
0.100	67.22	53.38	59.65	65.96	49.17	46.10	47.69	65.38	48.31	45.54	51.43	102.78	67.17	46.62	55.14	105.31
0.150	31.75	31.94	38.61	41.67	30.41	27.25	26.67	41.38	29.69	26.61	26.70	66.03	42.83	28.01	35.41	85.16
0.200	20.20	23.40	28.68	30.56	21.97	19.46	18.18	30.40	21.29	18.94	17.86	48.67	31.51	19.95	27.42	75.77
0.300	11.83	15.42	18.94	20.06	13.86	12.61	10.90	19.95	13.36	12.02	10.67	31.79	20.65	12.63	20.58	66.78
0.500	6.80	9.19	11.28	11.89	7.77	7.48	5.99	11.83	7.65	6.69	5.94	18.81	12.22	7.38	15.32	59.69
0.750	4.51	6.10	7.49	7.88	5.09	4.96	3.91	7.85	5.06	4.10	3.90	12.45	8.05	4.88	12.69	56.27
1.000	3.38	4.57	5.61	5.90	3.81	3.71	2.92	5.87	3.78	2.94	2.92	9.31	5.88	3.64	11.36	54.55

TABLE 8.2. SSANOS values for various control charts with $p_0 = 0.001$.

	Shewhart CCC-*r*			Bernoulli GLR				Bernoulli CUSUM				Bernoulli EWMA			Binomial GLR	
	r=2	*r*=3	*r*=4	p_{UB}=0.002	p_{UB}=0.005	p_{UB}=0.01	p_{UB}=0.02	p_1=0.002	p_1=0.005	p_1=0.01	p_1=0.02	λ=0.00001	λ=0.0001	λ=0.001	*n*=150	*n*=1000
p	*h*=122	*h*=439	*h*=901	*h*=4.39	*h*=6.19	*h*=7.03	*h*=7.84	*h*=4.35	*h*=6.05	*h*=6.80	*h*=7.63	*h*=0.0011	*h*=0.0017	*h*=0.0046	*h*=6.28	*h*=4.06
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]
0.0010	298031.0	299634.3	298309.9	297569.5	298960.0	298464.8	299735.1	299767.0	299362.4	295102.2	298457.0	297757.3	295233.9	300003.7	303609.1	279999.5
0.0015	92915.1	70208.1	56854.9	22478.6	29036.4	33194.0	37814.2	23974.2	44603.4	61788.6	81782.5	20586.6	22511.1	49642.4	30673.5	23326.5
0.0020	41082.4	26414.5	19429.5	8487.7	10083.8	11333.9	12848.2	8525.9	13829.0	21592.0	32847.3	9792.1	8356.8	15905.0	10714.9	9008.2
0.0025	22067.1	12927.2	9242.6	4968.4	5416.4	6014.5	6738.4	4948.0	6482.2	10111.2	16401.9	6409.4	4933.0	7417.0	5803.6	5336.5
0.0030	13376.4	7468.7	5369.7	3484.2	3522.9	3863.8	4297.2	3462.6	3876.2	5714.1	9439.5	4757.7	3480.7	4335.1	3788.4	3793.2
0.0040	6252.3	3379.1	2593.6	2174.0	1975.2	2104.4	2313.8	2161.4	2007.9	2590.7	4126.5	3146.2	2187.4	2146.3	2135.6	2469.8
0.0050	3535.1	1963.3	1633.6	1581.6	1338.9	1388.4	1509.0	1570.2	1329.2	1545.0	2288.0	2349.3	1593.3	1375.4	1451.7	1891.0
0.0060	2266.4	1325.0	1189.8	1242.3	1005.4	1018.8	1091.8	1233.5	990.0	1069.9	1472.4	1875.6	1254.4	1002.4	1092.8	1572.3
0.0070	1579.7	981.1	942.2	1024.1	802.8	795.0	844.7	1016.2	789.5	810.5	1046.9	1560.7	1034.5	784.5	876.2	1370.7
0.0080	1171.4	774.7	788.1	869.8	667.8	649.1	681.1	863.7	656.0	649.0	794.0	1336.3	879.8	644.7	732.5	1231.4
0.0090	910.5	639.9	680.5	756.9	573.1	546.2	568.8	750.6	562.2	541.1	634.4	1169.2	765.3	547.0	630.8	1135.0
0.0100	730.3	546.1	601.6	669.5	500.4	471.0	485.5	664.3	490.9	464.1	523.1	1038.3	677.1	474.9	554.5	1060.2
0.0125	473.8	403.3	470.5	518.4	382.6	349.8	351.5	515.0	374.6	343.0	361.2	812.1	526.6	356.9	431.4	935.1
0.0150	341.1	323.5	388.4	422.5	309.7	277.8	273.4	420.3	302.7	272.5	274.3	666.7	431.2	285.5	357.9	856.5
0.0200	215.0	235.7	288.3	308.9	225.2	197.7	187.2	307.6	218.2	194.4	184.3	491.8	317.2	203.3	276.6	761.9
0.0500	68.9	92.4	113.3	119.9	81.8	75.3	62.3	119.4	78.3	72.7	61.1	189.9	123.0	74.5	152.8	601.4
0.1000	34.0	45.9	56.4	59.5	38.6	37.3	29.5	59.3	38.1	34.1	29.4	93.9	60.6	36.8	113.4	549.8
0.2000	16.9	22.9	28.1	29.6	19.1	18.6	14.7	29.5	19.0	15.5	14.6	46.8	29.6	18.3	93.8	524.8
1.0000	3.4	4.6	5.6	5.9	3.8	3.7	2.9	5.9	3.8	2.9	2.9	9.3	5.9	3.6	78.4	504.4

TABLE 8.3. SSANOS values for various control charts with $p_0 = 0.1$.

	Shewhart CCC-r		Bernoulli GLR				Bernoulli CUSUM				Bernoulli EWMA			Binomial GLR	
	r=3	r=4	p_{UB}=0.15	p_{UB}=0.2	p_{UB}=0.3	p_{UB}=0.4	p_1=0.15	p_1=0.2	p_1=0.3	p_1=0.4	λ=0.001	λ=0.01	λ=0.1	n=10	n=97
p	h=5	h=10	h=4.14	h=5.22	h=6.19	h=6.73	h=4.10	h=5.13	h=6.03	h=6.54	h=0.1146	h=0.1715	h=0.4269	h=5.11	h=2.81
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
0.100	8115.25	4822.91	6138.55	6100.43	6061.25	6286.61	6121.76	6062.97	6037.83	6286.87	6168.22	6078.89	6042.41	6116.40	6238.07
0.110	5597.11	3126.11	1945.45	2245.56	2557.38	2824.51	2072.09	2568.20	3112.10	3556.03	1589.22	2426.33	3683.38	2404.67	1877.73
0.115	4706.68	2562.10	1251.83	1464.20	1698.73	1883.93	1349.31	1762.99	2305.97	2740.02	1030.58	1646.84	2939.35	1570.36	1227.08
0.120	3995.77	2116.95	871.30	1015.58	1184.95	1310.98	931.59	1253.67	1739.19	2136.84	738.25	1161.55	2367.15	1088.26	870.24
0.125	3399.37	1770.78	642.02	746.00	867.77	958.96	678.44	918.66	1330.92	1689.15	565.06	846.94	1921.23	786.32	658.01
0.130	2923.78	1488.23	497.55	571.98	663.53	732.22	517.41	694.15	1039.48	1350.97	454.65	641.35	1577.97	602.61	523.79
0.140	2193.21	1076.19	330.04	370.66	426.14	467.52	335.43	429.45	659.64	891.19	323.17	399.88	1092.83	398.26	365.47
0.150	1681.33	800.45	240.69	262.45	299.11	326.63	241.26	290.40	442.28	611.55	249.96	274.15	779.89	284.75	279.95
0.175	932.21	417.35	138.89	141.20	156.01	168.60	138.18	144.92	199.93	278.66	158.73	140.90	376.90	152.92	180.83
0.200	560.11	244.00	96.55	92.37	98.47	105.40	95.94	92.30	113.15	151.02	116.00	91.25	207.93	100.26	138.80
0.250	243.41	104.99	59.71	52.94	52.56	55.10	59.29	52.17	53.98	64.30	75.39	52.60	85.41	56.59	103.58
0.300	125.58	56.35	43.20	36.82	34.49	35.32	42.86	36.19	34.25	37.26	55.87	36.74	45.87	38.38	89.02
0.400	46.39	24.62	27.84	22.83	19.95	19.53	27.57	22.43	19.51	19.28	36.81	22.90	20.94	24.85	75.40
0.500	22.84	14.91	20.55	16.60	13.97	13.21	20.32	16.27	13.58	12.90	27.45	16.65	12.92	17.96	68.52
1.000	4.78	5.83	8.97	7.00	5.51	4.75	8.90	6.92	5.43	4.73	12.12	7.10	4.44	11.17	58.25

8.2 Approximating the Bernoulli GLR Chart by the Multiple Bernoulli CUSUM Charts

In Chapter 5, we studied the relationship between the binomial GLR chart and the binomial CUSUM charts, and found that the binomial GLR chart without a window is equivalent to a countable set of binomial CUSUM charts. The maximum value of the CUSUM tuning parameter p_1 to be included in the CUSUM set can be found by solving Equation (5.3) numerically.

For the Bernoulli case, a similar equivalence also exists between the Bernoulli GLR chart and the Bernoulli CUSUM chart. However, since now the estimate of p_1 in the Bernoulli GLR statistic is no longer the MLE but with an upper bound p_{UB} imposed, the maximum value of p_1 in the CUSUM set cannot be obtained by solving Equation (5.3) with setting the value of n on the right-hand side to be 1. Instead, we found that if no window is used, a Bernoulli GLR chart with an upper bound p_{UB} is equivalent to a countable set of Bernoulli CUSUM charts under the condition that the largest value of p_1, denoted by $p_{1\max}$, is equal to p_{UB}. If a window of size m_{B} is used, the exact equivalence no longer holds, but a Bernoulli GLR is still approximately equal to a set of Bernoulli CUSUM charts given that $p_{\mathrm{UB}} = p_{1\max}$.

This claim is confirmed by simulation. Based on the simulation results, we found that if the same control limit is used by both Bernoulli GLR and CUSUM charts, the Bernoulli GLR chart with an upper bound equal to p_{UB} performs almost identically to a combination of multiple Bernoulli CUSUM charts, given that the values of p_1 fall in the interval of $(p_0, p_{1\max}]$, where $p_{\mathrm{UB}} = p_{1\max}$. Three different levels of p_0 are considered, and the results are displayed in Tables 8.4-8.6, respectively. The number of individual CUSUM charts used in the combination is 1,000, and the p_1 values for each individual chart is selected in the same manner as in the binomial case, which is discussed in more detail in Section 5.4.

In Table 8.4, we assume $p_0 = 0.01$, and a shift occurs after the process is in steady state with $\tau = 10{,}000$. For $p_{\mathrm{UB}} = p_{1\max} = 0.02$, the GLR chart and the CUSUM combination have almost exactly the same in-control performance and steady state out-of-control performance at all shift sizes. When p_{UB} gets larger, for example, $p_{\mathrm{UB}} = 0.2$, the SSANOS value of the GLR chart is 7668.67 for detecting a shift as small as $p = 0.013$, while that value of the CUSUM chart is 7583.97. Although the two SSANOS values are still close, the small difference between

them is mainly due to the discreteness of the Bernoulli distribution, which seems to be more severe when p_{UB} or $p_{1\max}$ becomes relatively large. However, as the shift size p gets larger, say, $p \geq 0.02$, the difference between the two SSANOS values becomes negligible.

Similar conclusions can also be drawn from Tables 8.5 and 8.6 where $p_0 = 0.001$ and $p_0 = 0.1$, respectively. With $p_{\mathrm{UB}} = p_{1\max}$, the Bernoulli GLR chart can be well approximated by a set of 1,000 Bernoulli CUSUM charts even for very small shift sizes. Furthermore, we used 1,000 individual CUSUM charts in the combination here as an illustration. In fact it is not necessary to use such a large number of CUSUM charts for any practical application – using around 200 of them would generate SSANOS values that are sufficiently close to those of the GLR chart.

Since now we can easily use a set of Bernoulli CUSUM charts to represent the Bernoulli GLR chart, we could also apply the set of CUSUM charts to monitor the Bernoulli processes in practice. The advantage of applying such a combination over the GLR chart is that for some specific combination of ICANOS$\times p_0$ and p_{UB}/p_0, if h_{BGLR} cannot be directly obtained from Table 7.1 or Figure 7.1 and simulations are required to find the proper h_{BGLR}, it turns out that simulating a set of CUSUM charts is much faster than simulating a GLR chart. Also, using this approach avoids the need to specify the window size of the GLR chart, but it still requires that we use the same control limit as is used by the GLR chart. In addition, if a sufficiently large number of CUSUM charts are considered in the combination, its performance would be significantly improved compared to using only one CUSUM chart with $p_1 = p_{1\max}$. The combined Bernoulli CUSUM chart can be plotted in the same way as we plot the combined binomial CUSUM chart in Section 5.4. Although all Bernoulli CUSUM statistics are computed internally, only the maximum of them would actually be plotted, with the combined chart being very similar to a regular Bernoulli CUSUM or GLR chart (see Figure 6.2 in Section 6.3 as example plots of regular Bernoulli CUSUM and GLR charts). Thus, instead of using individual Bernoulli CUSUM chart to monitor the Bernoulli processes, we would recommend that either the Bernoulli GLR chart or a combination of multiple Bernoulli CUSUM charts being used for any practical application.

TABLE 8.4. Approximating the Bernoulli GLR charts with various upper bounds using a set of Bernoulli CUSUM charts when $p_0 = 0.01$.

	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo
	p_{UB}=0.02	p_{1max}=0.02	p_{UB}=0.05	p_{1max}=0.05	p_{UB}=0.1	p_{1max}=0.1	p_{UB}=0.2	p_{1max}=0.2
	h=4.38		*h*=6.19		*h*=7.01		*h*=7.84	
p								
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
0.010	29204.97	29258.19	29235.49	29131.30	28947.65	28889.92	29252.84	29259.87
0.013	4457.73	4457.97	5851.55	5862.43	6672.54	6645.49	7668.67	7583.97
0.015	2222.50	2220.62	2871.64	2873.62	3260.42	3260.09	3745.27	3741.24
0.020	836.64	836.71	995.56	996.86	1117.78	1117.65	1269.77	1270.18
0.025	489.97	490.02	534.43	535.19	593.13	592.00	667.17	667.05
0.030	343.62	343.56	348.16	348.06	380.46	380.35	424.79	425.20
0.040	214.55	214.54	194.45	194.51	207.21	207.15	228.31	228.59
0.050	155.80	155.93	131.63	131.83	136.49	136.37	148.94	148.96
0.060	122.47	122.50	98.88	98.79	100.00	99.89	107.78	107.78
0.070	100.91	100.90	78.91	78.87	77.91	77.81	83.12	83.10
0.080	85.84	85.79	65.68	65.67	63.58	63.53	67.08	67.11
0.090	74.63	74.56	56.22	56.25	53.46	53.41	55.88	55.85
0.100	65.96	65.94	49.17	49.20	46.10	46.10	47.69	47.54
0.150	41.67	41.68	30.41	30.39	27.25	27.21	26.67	26.66
0.200	30.56	30.54	21.97	21.96	19.46	19.45	18.18	18.16
0.300	20.06	20.03	13.86	13.87	12.61	12.61	10.90	10.89
0.500	11.89	11.89	7.77	7.78	7.48	7.48	5.99	5.99
0.750	7.88	7.88	5.09	5.09	4.96	4.96	3.91	3.91
1.000	5.90	5.89	3.81	3.81	3.71	3.71	2.92	2.92

TABLE 8.5. Approximating the Bernoulli GLR charts with various upper bounds using a set of Bernoulli CUSUM charts when $p_0 = 0.001$.

	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo
	p_{UB}=0.002	p_{1max}=0.002	p_{UB}=0.005	p_{1max}=0.005	p_{UB}=0.01	p_{1max}=0.01	p_{UB}=0.02	p_{1max}=0.02
	h=4.39		*h*=6.19		*h*=7.03		*h*=7.84	
p								
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
0.0010	297569.53	297917.68	298960.00	297959.30	298464.81	298535.36	299735.11	299481.07
0.0015	22478.56	22522.73	29036.40	29055.03	33194.04	33175.20	37814.15	37870.78
0.0020	8487.69	8489.25	10083.84	10081.79	11333.87	11356.71	12848.24	12836.19
0.0025	4968.39	4968.01	5416.39	5415.15	6014.49	6017.51	6738.40	6744.76
0.0030	3484.19	3484.94	3522.88	3525.67	3863.83	3858.85	4297.17	4296.10
0.0040	2174.01	2175.42	1975.15	1974.88	2104.40	2106.11	2313.82	2311.22
0.0050	1581.59	1580.46	1338.88	1338.29	1388.45	1389.46	1508.99	1509.65
0.0060	1242.34	1242.73	1005.38	1004.93	1018.80	1018.84	1091.80	1093.65
0.0070	1024.06	1023.75	802.80	803.74	794.96	794.70	844.69	844.11
0.0080	869.75	870.14	667.76	668.06	649.06	649.23	681.10	682.26
0.0090	756.94	757.13	573.09	572.52	546.16	546.44	568.76	568.38
0.0100	669.54	669.38	500.37	501.21	471.05	470.92	485.54	485.09
0.0125	518.38	518.88	382.62	382.49	349.76	349.78	351.50	351.61
0.0150	422.55	422.93	309.68	310.15	277.82	277.84	273.40	273.41
0.0200	308.94	309.31	225.22	225.19	197.70	198.02	187.19	187.30
0.0500	119.93	120.00	81.84	81.81	75.27	75.25	62.32	62.33
0.1000	59.45	59.48	38.58	38.60	37.34	37.32	29.49	29.52
0.2000	29.61	29.60	19.10	19.12	18.60	18.61	14.65	14.66
1.0000	5.90	5.90	3.81	3.81	3.71	3.71	2.92	2.92

TABLE 8.6. Approximating the Bernoulli GLR charts with various upper bounds using a set of Bernoulli CUSUM charts when $p_0 = 0.1$.

	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo	Bernoulli GLR Chart	Bernoulli CUSUM Combo
	p_{UB}=0.15 p_{1max}=0.15		p_{UB}=0.2 p_{1max}=0.2		p_{UB}=0.3 p_{1max}=0.3		p_{UB}=0.4 p_{1max}=0.4	
p	h=4.14		h=5.22		h=6.19		h=6.73	
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
0.100	6138.55	6130.64	6100.43	6112.80	6061.25	6064.40	6286.61	6291.99
0.110	1945.45	1944.70	2245.56	2237.98	2557.38	2543.87	2824.51	2802.56
0.115	1251.83	1251.72	1464.20	1461.95	1698.73	1691.75	1883.93	1872.89
0.120	871.30	870.85	1015.58	1015.67	1184.95	1182.93	1310.98	1308.91
0.125	642.02	643.49	746.00	746.73	867.77	868.99	958.96	959.27
0.130	497.55	497.53	571.98	571.89	663.53	663.83	732.22	732.37
0.140	330.04	330.53	370.66	370.46	426.14	426.38	467.52	468.25
0.150	240.69	240.76	262.45	262.76	299.11	299.36	326.63	326.96
0.175	138.89	138.86	141.20	141.04	156.01	155.98	168.60	168.54
0.200	96.55	96.58	92.37	92.38	98.47	98.44	105.40	105.34
0.250	59.71	59.66	52.94	52.94	52.56	52.64	55.10	55.12
0.300	43.20	43.21	36.82	36.78	34.49	34.50	35.32	35.33
0.400	27.84	27.84	22.83	22.85	19.95	19.93	19.53	19.53
0.500	20.55	20.53	16.60	16.57	13.97	13.98	13.21	13.21
0.750	12.41	12.40	9.83	9.82	7.93	7.93	7.21	7.21
1.000	8.97	8.96	7.00	6.99	5.51	5.51	4.75	4.75

Chapter 9

Conclusions and Future Research

9.1 Conclusions

When monitoring a process proportion using samples of size $n > 1$, the traditional Shewhart p- or np-charts are widely used because of their simplicity in design and application. Using only the most recent sample observation, a Shewhart chart is very powerful for detecting large shifts. However, since no past information is taking into account, the Shewhart chart is not effective in detecting small changes, which makes it incapable of having fast detection over a wide range of shift sizes. As an alternate approach, a CUSUM chart makes use of an additional tuning parameter so that it can be designed to have very quick detection of some specific shift sizes. Nevertheless, its performance is not as satisfactory if the actual size of the change is far away from what is expected. Another option is to use an EWMA chart that incorporates a smoothing parameter, through which the EWMA chart can be designed to be sensitive to either small or large shifts, but not all possible shift sizes at the same time. In addition, the introduction of an extra parameter in the CUSUM or EWMA chart makes the design of such a chart more complicated, and thus less appealing to practitioners.

For monitoring process proportions where samples are only obtained during certain sampling periods or when rational subgrouping is used to aggregate items into samples of size n, we propose using a new type of control chart based on the generalized likelihood ratio statistic that is calculated from a binomial random variable. The binomial GLR chart can be perceived as a combination of a set of countable number of binomial CUSUM charts each tuned to detect a particular shift size. In this sense, no tuning parameter is needed for the binomial GLR chart; instead, a maximum likelihood estimator for the shift size is used, which grants the GLR chart the ability to quickly detect different shift sizes over a wide range. This also means that in practice, the implementation of the binomial GLR chart could be easier than many other charts, because it requires specification of fewer control-chart parameters. In addition, the GLR chart

gives immediate estimates of both the process change-point and the shifted value of p, which could be helpful to process engineers in their post-signal process diagnosis.

Designing a binomial GLR chart requires specification of the size of a moving window as well as the control limit. Since calculating the GLR statistic requires maximization over all past samples, a moving window is used to relieve part of the computational burden. The choice of the window size is rather robust if a sufficiently large number of observations are included in this window. A rule of thumb is to select the window size so that it matches the target in-control ANSS value. In addition, it is found and confirmed by simulation that the control limit of the GLR chart is roughly linearly related to the natural log of the in-control ANSS, and this holds for many combinations of p_0 and n that are frequently encountered in practice. A linear model is provided in Equation (4.1) so that for a given ln(in-control ANSS) value, the control limit could easily be obtained using this model as long as $n \times p_0$ is not smaller than 0.1.

Performance of the binomial GLR chart is evaluated for various combinations of p_0 and n. Comparisons are made under the steady state scenario among the binomial GLR chart, the Shewhart-np chart, individual binomial CUSUM charts, the Shewhart-CUSUM combination, and the combination of multiple CUSUM charts. Results are obtained through simulation, which show that the binomial GLR chart is the most effective among all the charts in detecting a wide range of shift sizes, with its overall performance being at least as good as that of the others. In addition, the combination of multiple binomial CUSUM charts behaves quite similarly to the binomial GLR chart, turning it into another candidate for effective detection of multiple shift sizes.

The relationship between the binomial GLR chart and the binomial CUSUM chart is also investigated, and we find that the binomial GLR chart with a window could be well approximated by a countable set of binomial CUSUM charts. It turned out that if n is large, only a small number of CUSUM charts are needed to achieve a good approximation. However, more CUSUM charts are required if n is relatively small.

In summary, the binomial GLR chart has two major advantages over most of the other binomial control charts: it does not require users to specify values of numerous control chart

parameters, which greatly reduces the effort needed to design and implement this chart; also, it has very good overall performance for detecting a wide range of shift sizes.

To monitor Bernoulli processes with continuous inspection, the sample size n becomes 1, and it is more appropriate to apply a GLR chart that is directly based on the individual Bernoulli observations. Any charting scheme based on artificial subgrouping that generates samples of size $n > 1$ could not give an immediate signal if a large shift occurs, and thus is not desirable. However, the case of $n = 1$ is not just a simple special case of $n > 1$. When n is very small there is a non-negligible probability that $\hat{p}_1 = 1$, which results in an undefined likelihood ratio statistic. Handling this problem is not trivial, and several approaches are studied to overcome this. We choose to impose an upper bound p_{UB} on $\hat{p}_1$ to prevent it from being 1. This upper bound serves like an addition charting parameter of the Bernoulli GLR chart, which is similar to the tuning parameter p_1 of the CUSUM chart.

Designing a Bernoulli GLR chart requires specification of the upper bound p_{UB} in addition to the window size and control limit. The choice of window size follows the same rule of thumb as we found in the binomial case – it should be at least as large as the desired in-control ANOS value. Table 7.1 and Figure 7.1 are established to help in the selection of the control limit for different levels of p_0 ranging from 0.001 to 0.1. To choose an appropriate upper bound p_{UB}, one should use the same rule that people generally use to determine the value of p_1 for a CUSUM chart – a small value of p_{UB} is preferred if the interest is to detect small shifts, and a large value of p_{UB} is preferred if we expect large shifts to occur.

Comparisons among the Shewhart-type CCC-r chart, the Bernoulli EWMA chart and the Bernoulli GLR chart show that the Bernoulli GLR chart with an upper bound has better performance at most shift sizes. When compared to the Bernoulli CUSUM chart with $p_1 = p_{\mathrm{UB}}$, we find that the Bernoulli GLR chart is more effective in detecting small shifts such as $p < p_1$ (or p_{UB}); for any shift $p \geq p_1$ (or p_{UB}), the Bernoulli GLR chart is only slightly worse than the corresponding Bernoulli CUSUM chart.

Furthermore, we found that there also exists an approximate equivalence between the Bernoulli GLR chart with a window and an upper bound, and a countable set of the Bernoulli CUSUM charts, given that $p_{\mathrm{UB}} = p_{1\max}$, where $p_{1\max}$ is the largest value of p_1 in the

combination. In this case, the performance of the Bernoulli GLR chart can be easily approximated by a set of individual Bernoulli CUSUM charts. Usually a small number such as 200 CUSUM charts is sufficient to be used for a good approximation in practice.

It is well known that the Bernoulli CUSUM chart is very difficult to beat in terms of its overall steady-state performance for monitoring the Bernoulli processes. However, based on the simulation results, we recommend using the Bernoulli GLR chart over the CUSUM chart because, first of all, compared to the effort required for designing a Bernoulli CUSUM chart, it is easier to design a Bernoulli GLR chart based on the guidelines provided in Chapter 7. In addition, the Bernoulli GLR chart has much faster detection of small shifts than the Bernoulli CUSUM chart and almost the same performance for the other shifts.

9.2 Future Research Topics

Further research is needed to thoroughly understand the application of the GLR chart to monitor process proportions. In particular,

1. This dissertation has considered the binomial and Bernoulli GLR charts for detecting increases in p. Detecting decreases or both increases and decreases in p is of interest in some applications. Developments of such GLR charts are topics for future research.
2. To construct the Bernoulli GLR chart, here we rely on the maximum likelihood estimate of p_1 and impose an upper bound on it to prevent $\hat{p}_1 = 1$. There could be other types of estimate of p_1 that might not have the issue of $\hat{p}_1 = 1$. More work can be done to search for such estimates by considering the alternatives such as the EWMA estimate or Bayesian estimate.
3. In this dissertation, the GLR charts are all built based on the assumption that we are in the Phase II monitoring and the parameters such as p_0 are either known or can be estimated without error. This assumption is often violated in practical applications. Many works have been done to evaluate the effect of estimation error of the in-control parameter based on the Phase I data (see, for example, Lee et al. (2011) and

Zhang et al. (2011)). It is also important to study the performance of the binomial and Bernoulli GLR charts if p_0 is unknown and estimated using Phase I data.

4. An important application of the control charts based on binomial or Bernoulli observations is to monitor the high yield or near-zero nonconforming processes, in which case the number of nonconforming items could be very low such that it is measured by parts per million, and p_0 could be extremely small such as $p_0 < 0.001$. Choices of the values of the GLR parameters need to be further studied so that we can design a GLR chart that has good performance for monitoring this special situation.
5. The type of process change considered in this dissertation is assumed to be a sustained abrupt shift that presents in the process until being detected and the corresponding special cause being removed. Sometimes the process change occurs gradually or the shift might disappear after existing for some period of time. If this were the case, we need to evaluate the robustness of the GLR chart designed for detecting sustained shifts in the case of drifts or transient shifts.

Bibliography

Apley, D. W. and Shi, J. (1999). "The GLRT for Statistical Process Control of Autocorrelated Processes". *IIE Transactions*, 31, pp. 1123-1134.

Basseville, M. and Benveniste, A. (1983). "Design and Comparative Study of Some Sequential Jump Detection Algorithms in Digital Signals". *IEEE Transactions on Acoustics, Speech Signal Processing*, 31, pp. 521-535.

Basseville, M. and Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Applications*. Eaglewood Cliffs, NJ: Prentice-Hall.

Benveniste, A., Basseville, M. and Moustakides, G. (1987). "The Asymptotic Local Approach to Change Detection and Model Validation". *IEEE Transactions on Automatic Control*, 32, pp. 583-592.

Bennneyan, J. C. (2001). "Number-between *g*-type Statistical Quality Control Charts for Monitoring Adverse Events". *Health Care Management Science*, 4, pp. 301-318.

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.

Box, G. E. P., Jenkins, G. M. and MacGregor, J. F. (1974). "Some Recent Advances in Forecasting and Control". *Applied Statistics*, 23, pp. 158-179.

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

Calvin, T. W. (1983). "Quality Control Techniques for 'Zero Defects'". *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, CHMT-6, 3, pp. 323-328.

Cappizi, G. (2001). "Design of Change Detection Algorithms Based on the Generalized Likelihood Ratio Test". *Environmetrics*, 12, pp. 749-756.

Cappizi, G. and Masarotto, G. (2003). “An Adaptive Exponentially Weighted Moving Average Control Chart”. *Technometrics*, 45, pp. 199-207.

Chan, L. Y., Xie, M. and Goh, T. N. (1997). “Two-state Control Charts for High Yield Processes”. *International Journal of Reliability, Quality, and Safety Engineering*, 4, pp. 149-165.

Chan, L. Y., Lai, C. D., Xie, M. and Goh, T. N. (2003). “A Two-state Decision Procedure for Monitoring Processes with Low Fraction Nonconforming”. *Europe Journal of Operational Research*, 4, pp. 149-165.

Chang, T. C. and Gan, F. F. (2001). “Cumulative sum charts for high yield processes ”. *Statistica Sinica*, 11, pp. 791-805.

Cheng, C. S. and Chen, P. W. (2008). “The Design of Cumulative Count of Conforming Chart with Supplementary Runs Rules”. *Proceedings of the 9th Asia Pacific Industrial Engineering and Management Systems Conference*, pp. 204-209.

Crowder, S. V. (1987). “A Simple Method for Studying Run Length Distributions of Exponentially Weighted Moving Average Charts”. *Technometrics*, 29, pp. 401-407.

Crowder, S. V. and Hamilton, M. (1992). “An EWMA for Monitoring Standard Deviation”. *Journal of Quality Technology*, 24, pp. 12-21.

Dragalin, V. (1997). “The Design and Analysis of 2-CUSUM Procedures”. *Communications in Statistics-Simulation and Computation*, 26, pp. 67-81.

Ewan, W. D. (1963). “When and How to Use Cu-Sum Charts”. *Technometrics*, 5, pp. 1-22.

Gan, F. F. (1991). “An Optimal Design of Cusum Quality Control Charts”. *Journal of Quality Technology*, 23, pp. 279-286.

Gan, F. F. (1993). “An Optimal Design of CUSUM Control Charts for Binomial Counts”. *Journal of Applied Statistics*, 20, pp. 445-460.

Glushkovsky, E. A. (1994). “On-line g-control chart for attribute data”. *Quality and Reliability Engineering International*, 10, pp. 217-227.

Goh, T. N. (1987). “A Charting Technique for Control of Low-defective Production”. *Journal of Quality and Reliability Management*, 4, pp. 53-62.

Goh, T. N. (1987). “A Control Chart for Very High Yield Processes”. *Quality Assurance*, 13, pp. 18-22.

Gombay, E. (2000). “Sequential Change-Point Detection with Likelihood Ratios”. *Statistics & Probability Letters*, 49, pp. 195-204.

Han, D., Tsung, F., Hu, X. and Wang, K. (2007). “CUSUM and EWMA Multi-Charts for Detecting a Range of Mean Shifts”. *Statistica Sinica*, 17, pp. 1139-1164.

Hawkins, D. M. (1977). “Testing a Sequence of Observations for a Shift in Location”. *Journal of American Statistical Association*, 72, pp. 180-186.

Hawkins, D. M. (1981). “A New Test for Multivariate Normality and Homoscedasticity”. *Technometrics*, 23, pp. 105-110.

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

Hawkins, D. M., Qiu P. and Kang, C. W. (2003). “The Changepoint Model for Statistical Process Control”. *Journal of Quality Technology*, 35, pp. 355-366.

Hawkins, D. M. and Zamba, K. D. (2005). “Statistical Process Control for Shifts in Mean or Variance Using a Changepoint Formulation”. *Technometrics*, 47, pp. 164-173.

Hinkley, D. V. (1970). “Inference about the Change-Point in a Sequence of Random Variables”. *Biometrika*, 57, pp. 1-17.

Jiang, W., Shu, L. and Apley, D. W. (2008). “Adaptive CUSUM Procedures with EWMA-Based Shift Estimators”. *IIE Transactions*, 40, pp. 992-1003.

Kaminsky, F. C., Benneyan, J. C., Davis, R. D. and Bourke, R. J. (1992). “Statistical Control Charts Based on a Geometric Distribution”. *Journal of Quality Technology*, 24, pp. 63-69.

Kotani, T., Kuralmani, E. and Ohta, H. (2005). “Exponentially Weighted Moving Average Chart for High-yield Processes”. *Industrial Engineering and Management Systemsl of Quality Technology*, 4, pp. 75-81.

Kuralmani, V., Xie, M., Goh, T. N. and Gan, F. F. (2002). “A Conditional Decision Procedure for High Yield Processes”. *IIE Transactions*, 34, pp. 1002-1030.

Lai, C. D. and Govindaraju, K. (2008). “Reduction of Control-chart Signal Variability for High-quality Processes”. *Journal of Applied Statistics*, 35, pp. 671-679.

Lai, T. L. (1995). “Sequential Changepoint Detection in Quality Control and Dynamical Systems” (with discussion). *Journal of the Royal Statistical Society, Ser. B*, 57, pp. 613-658.

Lai, T. L. (1998). “Information Bounds and Quick Detection of Parameter Changes in Stochastic Systems”. *IEEE Transactions on Information Theory*, 44, pp. 2917-2929.

Lai, T. L. (2001). “Sequential Analysis: Some Classical Problems and New Challenges” (with discussion). *Statistica Sinica*, 11, pp. 303-408.

Lee, J., Wang, N., Xu, L., Schuh, A. and Woodall, W. H. (2011). “The Effect of Parameter Estimation on the Upper-Sided Bernoulli CUSUM Charts”. *Submitted for publication*.

Lorden, G. (1971). “Procedures for Reacting to a Change in Distribution”. *The Annals of Mathematical Statistics*, 42, pp. 1897-1908.

Lucas, J. M. (1976). “The Design and Use of V-Mask Control Schemes”. *Journal of Quality Technology*, 8, pp. 1-12.

Lucas, J. M. (1982). “Combined Shewhart-CUSUM Quality Control Schemes: A Robustness Study for CUSUM Quality Control Schemes”. *Journal of Quality Technology*, 14, pp. 51-59.

Lucas, J. M. and Saccucci, M. S. (1990). “Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements”. *Technometrics*, 32, pp. 1-12.

McCool, J. I. and Joyner-Motley, T. (1998). “Control Charts Applicable When the Fraction Nonconforming is Small”. *Journal of Quality Technology*, 30, pp. 240-247.

Montogomery, D. C. (2009). *Introduction to Statistical Quality Control, 7th Edition*. John Wiley & Sons, Inc., Hoboken, NJ.

Moustakides, G. V. (1986). "Optimal Stopping Times for Detecting Changes in Distributions". *The Annals of Statistics*, 14, pp. 1379-1387.

Nikiforov, I. V. (2001). "A Simple Change Detection Scheme". *Signal Processing*, 81, pp. 149-172.

Nishina, K. (1992). "A Comparison of Control Charts from the Viewpoint of Change-Point Estimation". *Quality and Reliability Engineering International*, 8, pp. 537-541.

Ohta, H., Kusukawa, E. and Rahim, A. (2001). "A CCC-r Chart for High-yield Processes". *Quality and Reliability Engineering International*, 17, pp. 439-446.

Page, E. S. (1954). "Continuous Inspection Schemes". *Biometrika*, 41, pp. 100-114.

Page, E. S. (1961). "Cumulative Sum Charts". *Technometrics*, 3, pp. 1-9.

Pignatiello, J. J., Jr. and Samuel, T. R. (2001). "Estimation of the Change Point of a Normal Process Mean in SPC Applications". *Journal of Quality Technology*, 33, pp. 82-95.

Reynolds, M. R., Jr. and Lou, J. (2010). "An Evaluation of a GLR Control Chart for Monitoring the Process Mean". *Journal of Quality Technology*, 42, pp. 287-310.

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, pp. 515-535.

Reynolds, M. R. Jr. and Stoumbos, Z. G. (2006). "Comparisons of Some Exponentially Weighted Moving Average Control Charts for Monitoring the Process Mean and Variance". *Technometrics*, 48, pp. 550-567.

Roberts, S. W. (1959). "Control Chart Tests Based on Geometric Moving Averages". *Technometrics*, 42, 1, Special 40th Anniversary Issue, pp. 97-101.

Rougée, A., Basseville, M., Benveniste, A. and Moustakides, G. (1987). "Optimum Robust Detection of Changes in the AR Part of a Multivariable ARMA Process". *IEEE Transactions on Automatic Control*, 32, pp. 1116-1120.

Runger, G. C. and Testik, M. C. (2003). "Control Charts for Monitoring Fault Signatures: Cuscore versus GLR". *Quality and Reliability Engineering International*, 19, 387-396. *Technometrics*, 1, pp. 239-250.

Samuel, T. R., Pignatiello, J. J., Jr. and Calvin, J. A. (1998a). "Identifying the Time of a Step Change with $\bar{X}$ Control Charts". *Quality Engineering*, 10, pp. 521-527.

Samuel, T. R., Pignatiello, J. J., Jr. and Calvin, J. A. (1998b). "Identifying the Time of a Step Change in a Normal Process Variance". *Quality Engineering*, 10, pp. 529-538.

Schwertman, N. C. (2005). "Designing Accurate Control Charts Based on the Geometric and Negative Binomial Distributions". *Quality and Reliability Engineering International*, 21, pp. 743-756.

Shewhart, W. A. (1931). *Economic control of quality of manufactured product*, Van Nostrand, New York.

Shu, L. J. and Jiang, W. (2006). "A Markov Chain Model for the Adaptive CUSUM Control Chart". *Journal of Quality Technology*, 38, pp. 135-147.

Siegmund, D. and Venkatraman, E. S. (1995). "Using the Generalized Likelihood Ratio Statistic for Sequential Detection of a Change-Point". *The Annals of Statistics*, 23, pp. 255-271.

Sparks, R. S. (2000). "CUSUM Charts for Signaling Varying Location Shifts". *Journal of Quality Technology*, 32, pp. 157-171.

Spliid, H. (2010). "An Exponentially Weighted Moving Average Control Chart for Bernoulli Data". *Quality and Reliability Engineering International*, 26, pp. 97-113.

Srivastava, M. S. (1994). "Comparison of CUSUM and EWMA Procedures for Detecting a Shift in the Mean or an Increase in the Variance". *Journal of Applied Statistical Science*, 1, pp. 445-468.

Sun, J. and Zhang, G. (2000). “Control Charts Based on the Number of Consective Conforming Items Between Two Successive Nonconforming Items for the Near Zero Nonconformaty Processes”. *Total Quality Management*, 11, pp. 235-250.

Szarka, J. L., III. and Woodall, W. H. (2011). “A Review and Perspective on Surveillance of Bernoulli Processes”. *Quality and Reliability Engineering International*, 27, pp. 735-752.

Szarka, J. L., III. and Woodall, W. H. (2011). “On the Equivalence of the Bernoulli and Geometric CUSUM Charts”. *To appear in Journal of Quality Technology*.

Weiβ, C. H. and Atzumüller, M. (2010). “EWMA Control Charts for Monitoring Binary Processes with Applications to Medical Diagnosis Data”. *Quality and Reliability Engineering International*, 26, pp. 795-805.

Westgard, J. O., Groth, T., Aronsson, T. and De Verdier, C. (1977). “Combined Shewhart-CUSUM Control Chart for Improved Quality Control in Clinical Chemistry”. *Clinical Chemistry*, 23, pp. 1881-1887.

Willsky, A. S. (1976). “A Survey of Design Methods for Failure Detection in Dynamic Systems”. *Automatica*, 12, pp. 601-611.

Willsky, A. S. and Jones, H. L. (1976). “A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems”. *IEEE Transactions on Automatic Control*, 21, pp. 108-112.

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

Woodall, W. H. and Adams, B. M. (1993). “The Statistical Design of CUSUM Charts”. *Quality Engineering*, 5, pp. 559-570.

Woodall, W. H. and Mahmoud, M. A. (2005). “The inertial properties of quality control charts”. *Technometrics*, 47, pp. 425-436.

Worsley, K. J. (1979). “On the Likelihood Ratio Test for a Shift in Location of Normal Populations”. *Journal of American Statistical Association*, 74, pp. 365-367.

Wu, Z., Jiao, J., Yang, M., Liu, Y. and Wang, Z. (2009). “An Enhanced Adaptive CUSUM Control Chart”. *IIE Transactions*, 41, pp. 642-653.

Wu, Z., Zhang, X. and Yeo, S. H. (2001). “Design of the Sum-of-conforming-run-length Control Charts”. *Europe Journal of Operational Research*, 132, pp. 187-196.

Wu, Z., Yang, M., Jiang, W. and Khoo, M. B. C. (2008). “Optimization Designs of the Shewhart-CUSUM Control Charts”. *Computational Statistics and Data Analysis*, 53, pp. 496-506.

Xie, M. and Goh, T. N. (1997). “The Use of Probability Limits for Process Control Based on Geometric Distribution”. *International Journal of Quality and Reliability Management*, 14, pp. 64-73.

Xie, M., Goh, T. N. and Kuralmina, V. (2002). *Statistical Models and Control Charts for High Quality Processes*. Kluwer, Boston, MA.

Xie, M., Lu, X. S., Goh, T. N. and Chan, L. Y. (1999). “A Quality Monitoring and Decision-making Scheme for Automated Production Process”. *International Journal of Quality and Reliability Management*, 16, pp. 148-157.

Xie, M., Tang, X. Y. and Goh, T. N. (2001). “On Economic Design of Cumulative Count of Conforming Chart”. *International Journal of Production Economics*, 72, pp. 89-97.

Xie, W., Xie, M. and Goh, T. N. (1995). “A Shewhart-like Charting Technique for High Yield Processes”. *Quality and Reliability Engineering International*, 11, pp. 189-196.

Yashchin, E. (1993). “Statistical Control Scheme: Methods, Applications and Generalizations”. *International Statistical Review*, 61, pp. 41-66.

Yeh, A. B., McGrath, R. N., Sembower, M. A. and Shen, Q. (2008). “EWMA Control Charts for Monitoring High-yield Processes Based on Non-transformed Observations”. *International Journal of Production Research*, 46, pp. 5679-5699.

Zhang, M., Peng, Y., Schuh, A., Megahed, F. M. and Woodall, W. H. (2011). “Geometric Charts with Estimated Parameters”. *To appear in Quality and Reliability Engineering International*.

Zhao, Y., Tsung, F. and Wang, Z. (2005). “Dual CUSUM Schemes for Detecting a Range of Mean Shifts”. *IIE Transactions*, 37, pp. 1047-1057.