On Development and Performance Evaluation of Some Biosurveillance Methods

Hongzhang Zheng

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

William H. Woodall
Sylvain S. DeLisle
Marion R. Reynolds, Jr.
Jeffrey B. Birch

July 5, 2011
Blacksburg, Virginia

Keywords: biosurveillance, control chart, CUSUM, detection delay, EARS, EWMA, false alarm rate, GLR, SARIMA, seasonality.

On Development and Performance Evaluation of Some Biosurveillance Methods

Hongzhang Zheng

Abstract

This study examines three applications of control charts used for monitoring syndromic data with different characteristics. The first part develops a seasonal autoregressive integrated moving average (SARIMA) based surveillance chart, and compares it with the CDC Early Aberration Reporting System (EARS) W2c method using both authentic and simulated data. After successfully removing the long-term trend and the seasonality involved in syndromic data, the performance of the SARIMA approach is shown to be better than the performance of the EARS method in terms of two key surveillance characteristics, the false alarm rate and the average time to detect the outbreaks.

In the second part, we propose a generalized likelihood ratio (GLR) control chart to detect a wide range of shifts in the mean of Poisson distributed biosurveillance data. The application of a sign function on the original GLR chart statistics leads to downward-sided, upward-sided, and two-sided GLR chart statistics in an unified framework. To facilitate the use of such charts in practice, we provide detailed guidance on developing and implementing the GLR chart. Under the steady-state framework, this study indicates that the overall GLR chart performance in detecting a range of shifts of interest is superior to the performance of traditional control charts including the EARS method, Shewhart charts, EWMA charts, and CUSUM charts.

There is often an excessive number of zeros involved in health care related data. Zero-inflated Poisson (ZIP) models are more appropriate than Poisson models to describe such data. The last part of the dissertation considers the GLR chart for ZIP data under a research framework similar to the second part. Because small sample sizes may influence the estimation of ZIP parameters, the efficiency of MLEs is investigated in depth, followed by suggestions for improvement. Numerical approaches to solving for the MLEs are discussed as well. Statistics for a set of GLR charts are derived, followed by modifications changing them from two-sided statistics to one-sided statistics. Although not a complete study of GLR charts for ZIP processes, due to limited time and resources, suggestions for future work are proposed at the end of this dissertation.

Acknowledgements

Dr. William H. Woodall and Dr. Sylvain S. DeLisle were instrumental in helping me design and carry out my research plans. I had many conversations with them as we went through the process of looking for and implementing research projects of statistical and practical significance. They watched as I made occasional mistakes and then guided me through the steps necessary to fix those errors. I thank them for their encouragement to openly share my thoughts and ideas and for their patience as I worked on this dissertation.

I owe many thanks to Dr. Marion R. Reynolds. His technical help and insightful suggestions made my work progress more smoothly. I also would like to thank Dr. Jeffrey B. Birch for all the ideas he shared with me. Along with Dr. Woodall, he gave me the encouragement that kept me moving forward during difficult times, for which I shall always be grateful.

A special thank goes to Dr. Holly D. Gaff at Old Dominion University for her help with the disease outbreak simulation and to Dr. Jinfeng Wei at Maryville University for some technical discussions.

I am forever grateful to my mother for her exemplary dedication and inspiration during my study and work in China and the U.S.. There really aren't words that can express the depth of gratitude I feel for my wife, Meiying. Her strength and encouragement have never faltered. She and our two wonderful children (Merrick, Timothy) have been my greatest joy and consolation during this dissertation research.

Contents

List of Figures

List of Tables

Part I

A Comparison of SARIMA and the EARS W2 Methods for Biosurveillance

Chapter 1

Introduction

Syndromic surveillance is the monitoring of health-related data sources using pre-diagnostic medical or other information to detect the release of a bioterrorism agent or the occurrence of a natural disease outbreak [1]. Many public health agencies now have real-time access to a large volume of daily data from multiple sources, such as numbers of emergency room visits, over-the-counter drug sales, school absentee rates, and numbers of doctors' office visits.

In general, health-related surveillance systems involve the following stages: data preparation, count prediction, aberrancy detection, and outbreak signal evaluation. In the past decade, researchers and public health practitioners have put much effort into developing automated surveillance systems to detect disease outbreaks rapidly with false alarm rates acceptable to the public health infrastructure [2, 3, 4]. Much of the work has been on developing new forecasting algorithms to predict counts and developing new methods to detect any aberrancy [1, 3, 5, 6].

The BioSense program of the Centers for Disease Control and Prevention (CDC) designed a surveillance system, the early aberration reporting system (EARS), which involved a 'C' family of detection algorithms (C1, C2 and C3) and W2c methods (modified

versions of C2) [7]. The EARS has been widely used in the United Sates and abroad, due to its simplicity of implementation and the limited amount of baseline data required.

The seasonal autoregressive integrated moving average (SARIMA) model of Box, Jenkins, and Reinsel [8] has been applied to fit time series data in many fields either to better understand the underlying process or to predict future points of the series. A well-developed SARIMA model takes into account historical fluctuations, trends, seasonality, cycles, prediction errors, and nonstationarity of the time series [9].

The objective of our research was to develop a surveillance system based on the SARIMA model and to compare its performance to the EARS W2 method in two different scenarios, (i) using authentic surveillance data and simulated outbreak data derived from a specified epidemic model; and (ii) using simulated surveillance data and simulated outbreak data provided by the CDC. We expected that the SARIMA approach would be able to detect outbreaks more quickly.

In particular, we describe our modeling approach to detecting aberrations and apply this method to a disease report series from Baltimore, Maryland and to the CDC data. We compare the surveillance methods on the basis of their sensitivity, specificity, and average times to outbreak detection given specified false alarm rates. Finally, we discuss the issues and challenges in the implementation of the SARIMA method in practice.

Chapter 2

Methods

2.1 Data preparation

Use of simulated outbreaks superimposed onto authentic data can be very informative when we evaluate syndromic surveillance algorithms. Simulated outbreaks based on authentic syndromic information enable one to perform sensitivity analyses and to make the monitoring algorithms more efficient [2, 10]. Surveillance systems, however, need to be tested rigorously under a range of conditions in realistic settings. Our outbreaks were simulated using a mathematical model, which was developed to be able to create a virtual epidemic whose characteristics could be used to test the sensitivity and specificity of syndromic surveillance systems for detecting influenza like illnesses (ILI).

The mathematical equations for generating the epidemic used in our outbreak simulations follow the classic epidemic structure as discussed in [11]. We have the following system of equations:

$$
\begin{aligned}
\frac{\partial X}{\partial a} + \frac{\partial Y}{\partial t} &= -(\lambda(a,t) + \mu(a))X(a,t) \\
\frac{\partial H}{\partial a} + \frac{\partial H}{\partial t} &= \lambda(a,t)X(a,t) - (\delta + \mu(a))H(a,t) \\
\frac{\partial Y}{\partial a} + \frac{\partial Y}{\partial t} &= \delta H(a,t) - (\gamma + \alpha(a) + \mu(a))Y(a,t) \\
\frac{\partial Z}{\partial a} + \frac{\partial Z}{\partial t} &= \gamma Y(a,t) - \mu(a,t)Z,
\end{aligned}
\tag{2.1}
$$

where X is the number of susceptible people, H is the number of people who are infected but not yet infectious, Y is the number of infectious people, Z is the number of people who are recovered and immune, a is the age class, t is the time period, λ is the per capita force of infection, μ is the fraction in each age class of infectious individuals that die as a direct result of the infection, δ is the parameter associated with the latency period, α is the fraction in each age class of infectious individuals that die as a direct result of the infection, and γ is the fraction of infectious individuals that moves into the recovered/immune class. The constants required to specify the model were selected to reflect the population of Baltimore, MD.

The simulated outbreak data provided benefits such as:

1. The characteristics of signals are similar to those seen during influenza outbreaks;
2. The susceptible population has a size and age structure that is similar to that of Baltimore, and reflects recent age-specific birth and death rates in susceptible populations;
3. The latency period and infectious period are similar to those observed with influenza; whereas the case mortality was adjusted to mimic severe influenza;
4. The prevalence of infectious individuals changes over time;

5. The outbreaks reflect a day-of-week (DOW) effect.

The simulated outbreaks generated by this system are also informative for spatiotemporal surveillance research, but we aggregated the spatial, age, and severity-structured information to obtain temporal data. The simulated outbreak derived from the specified epidemic model with the origin date October 2, 2002 is displayed in Figure 2.1.

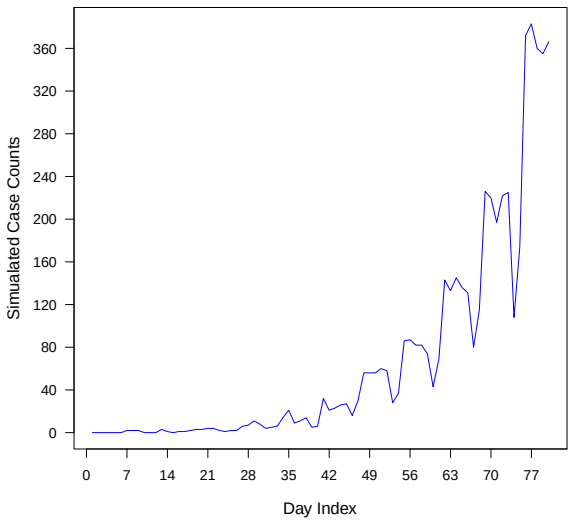


Figure 2.1: Epidemic model based outbreak simulation data for origin injection October 2, 2002

In terms of the data sources, there were three types of data for this research, i.e., background data, model-based simulated outbreak data (MBSD) and the combined data. Background data were the case detection algorithm (CDA) specified authentic syndromic counts obtained by applying the BioSense_M1, or the AndNegEx_M2 CDA to the Veterans' Affairs Maryland Healthcare System (VAMHCS) institutional historical databases; combined data were the summation of CDA-specific background data with the simulated data. The MBSD were simulated outbreak signals generated from the epidemic model. MBSD casecounts were multiplied by the case-finding sensitivity of a CDA before being added to a CDA-specific background casecounts to reflect the fact that the not all cases

would be identified by the CDA. More specifically, simulated data for BioSense_M1 and AndNegEx_M2 were obtained by multiplying the MBSD by 0.63 and 0.69 respectively.

The same MBSD was used in each outbreak considered in our study. For the purpose of the comparison of surveillance algorithm performance, the duration of each outbreak was set at 80 days, which would make the disease outbreak detectable in each injection cycle. When superimposing the simulated outbreak data on the background data, we carefully matched the DOW effects of the simulated data and the background data. We implemented 52 injection cycles in our study with the first outbreak starting on the first day of the time period studied. For any two adjacent injection cycles, the gap between their origin days was set at seven days.

2.2 Count prediction

2.2.1 EARS W2c

The CDC BioSense program developed the early aberration reporting system (EARS), which was intended to detect increases in counts or rates with health-related data. EARS has been used as a standard surveillance system in the United States and abroad since September 11, 2001 [12].

Originally, there were three detection methods used within EARS, referred to as C1, C2 and C3. The C2 method uses the counts from seven baseline days with a 2-day lag from the current day to calculate the sample average and sample standard deviation. Due to its simplicity and ease of implementation, the C2 method has been widely used in various cities, counties, and state public health officials in the United States and abroad [3, 12].

A modified version of C2, called W2c, is similar to C2, but it separates the weekdays and weekends/holidays to calculate the sample average and sample standard deviation. The study on the EARS algorithms indicates that compared to the C2 method, the W2c method can detect outbreaks rapidly [3].

The observed case counts, e.g., the numbers of visits at a particular hospital emergency department with a specific syndrome, are separated into weekdays and weekends/holidays first, and then the counts of weekdays and weekends/holidays are treated as different time series data with separate day indices.

The case of weekdays is taken as an example to illustrate how the W2c statistics are calculated. We let t be the time series index, and $y(t)$ be the observed case counts for weekday t. The W2c statistic, a residual divided by its standard deviation, is expressed as

$$W2c(t) = \frac{y(t) - \overline{y}(t)}{S(t)}, \tag{2.2}$$

where $\overline{y}(t)$ and $S(t)$ are the moving sample mean and standard deviation calculated with a 2-day lag respectively, i.e.,

$$\overline{y}(t) = \frac{1}{7} \sum_{i=t-3}^{t-9} y(i) \tag{2.3}$$

and

$$S(t) = \sqrt{\frac{1}{6} \sum_{i=t-3}^{t-9} [y(i) - \overline{y}(t)]^2}. \tag{2.4}$$

The value of is replaced by 1 if $S(t) < 1$.

The W2c method signals if the statistic in Equation (2.2) exceeds a given threshold. The thresholds such that the resulting value of the recurrence interval (RI) is a specified value can be found in the reference table provided by CDC BioSense [13]. For instance, for a threshold of 2.4, it indicates that the observed count would be 2.4 or more standard

deviations above the 7- day baseline sample mean on average once for every fixed 30 day period, i.e., the recurrence interval corresponding to the threshold value of 2.4 is 30 days. In our comparisons between EARS W2c method and the SARIMA approach, we used other metrics such as false alarm probabilities and average detection times instead of the recurrence interval, since we are primarily interested in how quickly the surveillance methods detect an outbreak.

2.2.2 SARIMA

Preliminary analysis

The electronic medical record database of VAMHCS in the Baltimore area contains records on disease incidence counts from January 2, 1999 to December 31, 2006, with eight years of data in total. The information in the database includes International Classification of Diseases, Ninth Revision (ICD-9) codes from the emergency room 'chief complaint' data. We generated several time plots for BioSense_M1 and AndNegEx_M2 disease case count series. We analyzed these to better understand the historical data and to guide the Box-Jenkins SARIMA modeling. These graphs included the following: (i) number of recorded daily cases of an 8 year period to visually portray the entire series of data; (ii) number of recorded daily cases for 2005 and 2006 to more carefully depict recent patterns; (iii) number of recorded daily cases of some individual months to explore the day-of-week effect; (iv) and plots of weekday and weekend data separately to determine if separate SARIMA modeling was needed based on this factor. The daily counts determined by BioSense_M1 are plotted in Figure 2.2 to illustrate some data features. The points on the horizontal dashed line in Figure 2.2 are calculated by the Lowess smother which uses locally-weighted polynomial regression [14].

From Figure 2.2, one can note a trend toward an increase in case counts of the Lowess smoothing, with the variation becoming larger over the time. Annual seasonal weather variations cause characteristic cyclic series features. Specifically, fewer counts appear generally around the middle of years while peaks occur around the months of November, December and January. To look more closely at the nature of the daily time series, plots of some months were also constructed. These time series plots displayed a few general predictable patterns, such as characteristic day-of-week (DOW) behavior. Specifically, daily counts are generally lower on weekends and higher early in the work week as displayed in Figure 2.3. The autocorrelation of syndromic surveillance data is another common issue. For seasonal diseases, like the common cold, the number of emergency department visits on one day will be correlated with counts on previous days [10].

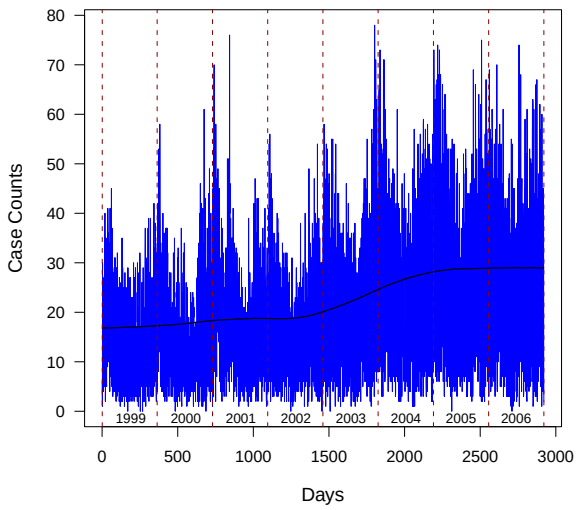


Figure 2.2: Case counts of influenza-like-illness (ILI) visits to the veteran affairs health care facilities in the Baltimore area. Cases are identified by the BioSense_M1 case detection algorithm.

Each of our incidence time series displays various characteristics, such as trends, seasonality, a DOW effect, and substantial autocorrelation. We used the Box-Jenkins SARIMA modeling method on the count data because it is designed to properly account

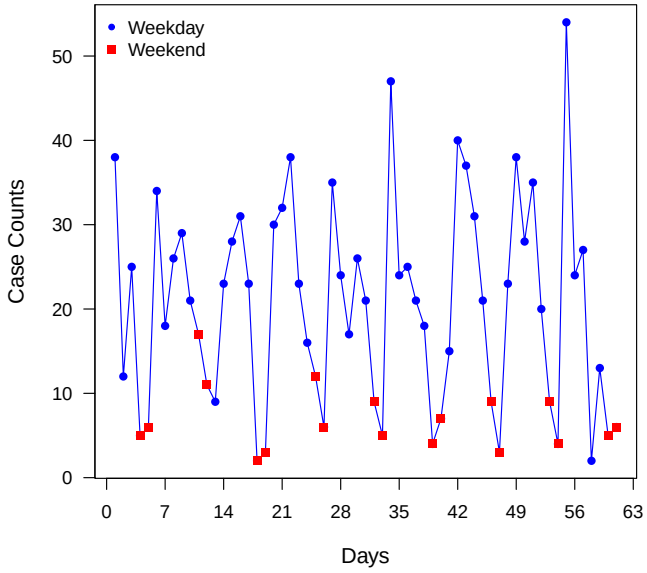


Figure 2.3: Case counts ranging from Wednesday, October 2, 2002 to Tuesday, December 3, 2002 showing day-of-week effect. Blue dots indicate weekdays, red dots indicate weekends.

for these issues [8]. If a suitable SARIMA time series model fits the data, it can be used to generate one-step-ahead minimum variance forecasts of the counts [15]. Then the forecast errors, the differences between the observations and the model forecasts, should be approximately independently and identically distributed (i.i.d.) when there are no outbreaks.

If we attempt to find a SARIMA model to fit the whole data sequence (i.e. without separation of weekdays and weekends/holidays), the standard error of the forecasts will be relatively larger, in part because of high counts on weekdays and low counts on weekends/holidays. The predicted case counts based on the resulting model will be consequently less precise. Just as the CDC C2 method was modified to W2c, we decided to divide the case counts data into weekdays and weekends/holidays, and found the corresponding appropriate SARIMA models. Although the SARIMA approach could be

used to model the weekly periodicity, it could not be expected to perform well with the less-predictable federal holidays. We therefore felt that the performance of the SARIMA models could also be enhanced by modeling the weekdays and the weekends/holidays time series separately.

There are ten VAMHCS holidays each year, which are the U.S. federal holidays defined by the U.S. Office of Personnel Management. The exploration of background data indicated the nature of patient visits to hospitals on holidays was very similar to visits on Saturday or Sunday, and most of the holidays occurred on weekdays. To improve the precision of the SARIMA forecasts for weekdays, in our study, the case counts of holidays occurring on Monday through Friday were replaced by the average counts of the most recent past three weekdays.

Modeling data

The case counts illustrated in Figure 2.2 indicate that the variation is not homogenous over the eight year period, and the variation becomes larger after 2003. Figure 2.2 also shows that there are some extreme counts (i.e., outliers) in the data series. In order to improve the reliability of SARIMA models and precision of SARIMA forecasts, some of these values are replaced by the averages of the past three daily counts. To compare the surveillance performance of EARS W2c to the SARIMA approach in different circumstances, we picked two contiguous years 2002-2004 in our study. The year of 2002-2003 was representative of a low-count year, whereas the year 2003-2004 was representative of a high-count year. Outbreaks were injected into these years with the two prior years in each case used as a baseline to fit the initial SARIMA models. Fifty-two separate injections were made with the initial injection starting dates separated at weekly intervals.

The combination of two CDAs (BioSense_M1, AndNegEx_M2) and two time periods (02-03, 03-04) generated four studied data streams, referred to as BioSense_M1_0203, BioSense_M1_0304, AndNegEx_0203, and AndNegEx_0304. The comparison of disease surveillance methods in our paper is based on these four data streams along with two data sets generated by the CDC.

Development of SARIMA model

Our initial explorations indicated that the data series studied were not stationary in both mean and variance. A square root transformation was applied to make them stationary, and thus this assumption of the Box-Jenkins modeling became satisfied [9].

Once transformations were used to make the data series stationary in both mean and variance, we used Box-Jenkins techniques to develop forecasting models [9, 16]. The Box-Jenkins approach to time series forecasting with seasonality can be implemented using an appropriate seasonal ARIMA model represented by SARIMA$(p,d,q)(P,D,Q)_s$ described as follows:

$$\Phi_P^*(B^s)\Phi_p(B)(1-B)^d(1-B^s)^D y_t = \Theta_Q^*(B^s)\Theta_q(B)\varepsilon_t, \tag{2.5}$$

where B is the backshift operator, d is the order of the differencing of the non-seasonal component, D is the order of the seasonal differencing, y_t is the observation at time t, ε_t is a sequence of independently distributed random shocks assumed to be normally distributed with a mean of zero and constant variance σ^2. In addition, $\Phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$ is an autoregressive polynomial of order p for non-seasonal component; $\Theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^p)$ is moving average polynomial of order q for non-seasonal component; $\Phi_P^*(B) = (1 - \phi_1^* B^s - \phi_2^* B^{2s} - \cdots - \phi_P^* B^{Ps})$is an autoregressive polynomial of order P for seasonal component; and $\Theta_Q^*(B) = (1 - \theta_1^* B^s -$

$\theta_2^* B^{2s} - \cdots - \theta_Q^* B^{Qs}$) is a moving average polynomial of order Q for seasonal component.

The SARIMA model development for to detect outbreaks in the first year was based on data from January 2, 1999 to August 13, 2002; data after this period were used to evaluate the surveillance methods. In order to develop a suitable SARIMA model for each series, various identification, estimation, and model evaluation steps were implemented. We checked the sample autocorrelation functions (ACF) and partial autocorrelation functions (PACF) to tentatively identify the most appropriate models. Model parameters were estimated.

We checked the adequacy of the SARIMA models by examining the residuals and forecast errors. Model adequacy for the estimation period was quantified based on the sum of squared residuals, and model forecast performance of the forecast periods was examined using mean absolute deviation of errors, the root mean squared error, Ljung-Box values, and QQ plots.

SARIMA model forecast

There are three types of input data streams in our study in terms of the data sources: background data, injected data, and combined data. Background data are the authentic case counts from the VAMHCS database identified by the particular CDA; injected data are the simulated case counts generated by the epidemic model for VAMHCS; and the combined data are summation of the background data and injected data. Background data and combined data were divided into two stages in our study: baseline data, which are the data before the outbreak was injected, and prospective data which are data after the outbreak has begun.

To improve the precision of the SARIMA forecast, the forecast horizon was always

set to 1 (i.e. one-step-ahead forecast). The most current information was consequently incorporated into the prediction procedure. In other words, in each injection cycle, the most current injected data along with the background data (i.e., combined data) are used to forecast the future case counts of this cycle.

Standardized Z-scores of SARIMA (SZS), the differences in the observed values and the SARIMA forecasts divided by the corresponding forecast standard errors, provides the measure of how far a current observation is from the non-outbreak situation. Large values of these Z-scores result in a signal that an outbreak exists for the SARIMA approach.

2.3 Aberrancy detection

The temporal aberrancy detection algorithms of syndromic surveillance are algorithms that sequentially evaluate the departure of the observed count or rate of health events from what would be expected based on previous experience when there is no outbreak [17].

In our project, the surveillance statistic values of the EARS and SARIMA methods are the W2c and SZS values, respectively. We determined the outbreak signals by varying some upper thresholds for both the EARS and the SARIMA methods. For a specific threshold on a given day, if the surveillance statistic value of the background data was below the specified threshold and the statistic value for the combined data was above the threshold, this signal was called a true positive. If the surveillance statistic values of both background data and combined data were above the specified threshold, the signal was called a false alarm (i.e., false positive). Figure 2.4 demonstrates the determination of

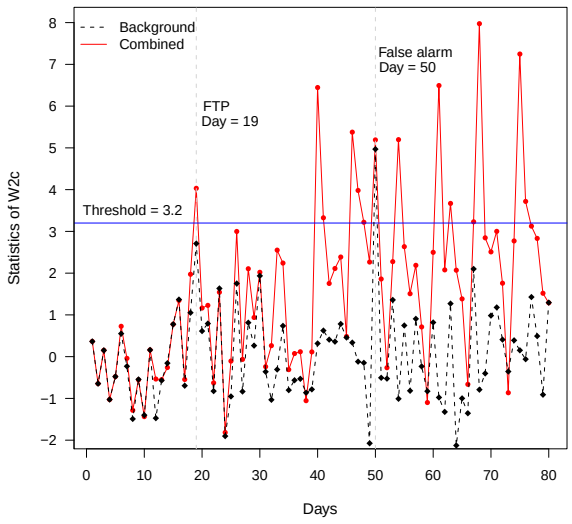


Figure 2.4: Illustration of true positives and false alarm

true positives and a false alarm for the threshold value of 3.2 in the 23th injection cycle of BioSense_M1_0203 by W2c method. There was one false alarm occurring on day 50, and the first true positive appeared on day 19. In this cycle, the detection delay was 18 days.

2.4 Surveillance evaluation

We programmed software that simulated a surveillance system as if it was performing prospectively on historical authentic datasets. The outbreak injection-aberrancy detection cycle begins when factitious cases originating from the mathematical model outlined in section 1.4 are injected into the background counts obtained by applying a particular CDA to historical relational databases developed at the VAMHCS. To identify the injected outbreak, we used either the EARS or the SARIMA aberrancy-detection methods. Starting from the injection day, the aberrancy detection statistic was prospectively applied daily once using the combined time series, and once more using the background-only time

series. On any given day, a statistical alarm issued in the background-only time series was considered a false-positive alarm. Conversely, an alarm issued in the combined dataset that had no counterpart in the background-only dataset constitutes a true-positive alarm. The cycle ends after 80 days, at time period that we had found sufficient to issue at least one true-positive alarm at the least sensitive threshold. To increase the robustness of our findings, the outbreak injection-aberrancy-detection cycle was repeated 52 times, each time shifting the timing of the injection to a different week of the study period.

The main metrics used in biosurveillance to evaluate an outbreak detection method is the timeliness. To compare two surveillance approaches, they need to be equivalent under some performance metric when there is no outbreak [6]. We adopted the activity monitoring operating characteristic (AMOC) curve [2], which is a plot of the average delay to the first true positive against the false alarm rate. An AMOC curve thus associates a timeliness score to each false positive rate of interest. For our scores, we used the average detection delay based on our 52 injection simulations. We calculated this delay for some false alarm rates (FAR) selected to be practical from the public health surveillance perspective.

The false alarm rate is the probability of an indication that an outbreak occurs in the background data investigated in which we assumed there were no outbreaks. Other metrics, such as the average time between signal events, could have been used as well.

In our study, we set the duration of the outbreak injection in each injection cycle at eighty days, which led to each injection cycle producing at least one true positive at the least sensitive threshold tested. The average delay used for AMOC curves was thus calculated based on all 52 injection cycles, where the separate injections were made at weekly intervals.

The model-based outbreak simulation data was generated using Matlab R2008a. The SARIMA modeling and forecasting were implemented by SAS 9.1. The aberrancy detection and surveillance performance evaluation were accomplished using R 2.10.1.

Chapter 3

Results

3.1 SARIMA model

As mentioned in section 2, the background case counts for each CDA were divided into weekday and weekend time series, based on which we found the most suitable SARIMA models for forecasting.

The annual period used by surveillance researchers is usually 364, or 365, or 365.25 [5, 15, 18, 19]; 364 is preferable because it will allow for DOW effects and often shows a greater degree of correlation than others [5]. In our study, we treated the annual seasonal period as 364 days, 52 weeks in one year; the cyclic periods of weekdays and weekends were of length 260 and 104, respectively. Since the data series we studied in this project were the combinations of two CDAs (BioSense_M1, AndNegEx_M2), two DOW categories (weekdays and weekends), there were subsequently four time series used for modeling and forecasting.

Examination of the data series indicated that they were non-stationary in the mean and variance. As shown in Figure 2.2, there is a slight trend in the mean, and some fluctuation in variance. To achieve the stationary requirements for implementing Box-Jenkins

ARIMA models, these data series required transformations and/or differencing. We conducted differencing to account for non-stationary means, and employed logarithmic and square root transformations to account for unstable variances. When using the logarithmic transformation, all explored SARIMA models lacked the seasonal component to account for the annual cycles of case counts, which was one of the apparent characteristics in the time series plot of the background case counts. This missing representation might have resulted from the fact that logarithm-transformed data counts were not large enough to catch the seasonal information for the Box-Jenkins SARIMA modeling techniques. For example, the mean of the background counts of the AndNegEx_M2 weekend baseline was 4.92, and its corresponding value was 0.692 after the logarithmic transformation.

Following the classical procedures of Box-Jenkins SARIMA modeling [16] (i.e., model identification, parameters estimation, and model evaluation), we were able to meet the stationary assumptions and develop adequate SARIMA models for all time series with a square root transformation. As an example of SARIMA modeling and analysis, we focus on the specific case BioSense weekday (with the holiday counts modification described in section 2) in year 2002-2003 (i.e., BioSense_wkday_0203). For the case of BioSense_wkday_0203, the baseline data used for the SARIMA modeling spanned from January 4, 1999 to August 13, 2002, while the rest of the data were used for the model validation. In the stage of model identification, autocorrelograms (ACF, PACF) were plotted to explore the characteristics of the time series, and to show the possible orders of the components for the SARIMA models. Some absolute coefficients of ACF and PACF at lags of 260 and multiples of 260 were statistically significant, which indicated that seasonal components should be included in the ARIMA model. Various potential forms of the ARIMA model with seasonality were explored interactively to search for the most appropriate model for each series. In the estimation stage, the coefficients of SARIMA models were calculated and statistical significance of each coefficient of the SARIMA model was eval-

uated. In the forecast stage, one-step-ahead forecasting was conducted to incorporate the most current surveillance data; forecasts and forecast errors were calculated to evaluate the performance of SARIMA model.

The most adequate SARIMA model should meet the model search criteria including a smaller Akaike's Information Criterion (AIC) value, a minimum mean absolute deviation (MAD), a minimum root mean square root error (RMSE), normality of residuals, and overall simplicity of model. Based on evaluation of statistical fit and model forecasting capabilities, the final model selected for this series was in the form of SARIMA (1,0,1)(0,1,1)260. Some diagnostic statistics (e.g., residuals, MAD, RMSE and Ljung-Box statistics) were calculated and related graphs (e.g., ACF, PACF and Q-Q) were plotted. The ACF, APCF and Q-Q plots of the residuals of the model indicated the resulting model provided a good fit to this data.

For example, using weekday data before August 16, 2004 as the baseline data, the statistical model for count forecasting on August 16, 2004 was

$$(1 - 0.998B)(1 - B^{260})y_t = (1 - 0.959B)(1 - 0.498B^{260})\varepsilon_t. \tag{3.1}$$

As mentioned earlier, we updated our forecast each day in the outbreak detection algorithm. When move forward the baseline data, a series of forecasts as well as forecast errors were generated.

Although developing models for all different data series (combinations of CDA and day of the week separation) separately, final models were shown in all of the same form, i.e., SARIMA$(1,0,1)(0,1,1)_s$. More specifically, the ARIMA models for weekday counts were all of the form SARIMA$(1,0,1)(0,1,1)_{260}$ and models for weekend counts were all of the form SARIMA$(1,0,1)(0,1,1)_{104}$. The coefficients of the SARIMA models change with

the time period since the models are refit each day. The coefficients for the four different models are given in Table 3.1, where the forecasting dates range from August 14, 2002 to August 18, 2002.

Table 3.1: Estimated coefficients of SARIMA(1,0,1)(0,1,1)s models for forecasting counts from 08/14/2002 - 08/18/2002

CDA	DOW	Forecasting Date	AR(1)	MA(1)	SMA(1)
BioSense	Wednesday	8/14/2002	0.929	0.785	0.543
BioSense	Thursday	8/15/2002	0.927	0.784	0.545
BioSense	Friday	8/16/2002	0.928	0.782	0.550
BioSense	Saturday	8/17/2002	0.834	0.639	0.699
BioSense	Sunday	8/18/2002	0.835	0.641	0.699
AndNegEx	Wednesday	8/14/2002	0.986	0.911	0.433
AndNegEx	Thursday	8/15/2002	0.986	0.910	0.438
AndNegEx	Friday	8/16/2002	0.985	0.909	0.431
AndNegEx	Saturday	8/17/2002	0.999	0.928	0.681
AndNegEx	Sunday	8/18/2002	0.997	0.925	0.685

3.2 Surveillance performance

AMOC curves (average delay versus false alarm rate) were plotted to compare the timeliness of the EARS and SARIMA methods as shown in Figure 3.1. Numbers of false alarms were determined by the specified thresholds and the numbers of false positives in the background data studied; false alarm rates were the ratios of numbers of false alarms to the number of days in the background data used. In order to make the comparison meaningful, false alarm rates were not allowed to exceed 0.1, which includes values that would be used by practitioners. Panels A, B, C and D of Figure 3.1 represent the AMOC for the four data series, AndNegEx_M2_0203, AndNegEx_M2_0304, BioSense_M1_0203, and BioSense_M1_0304, respectively. Average delays of the SARIMA approach are consistently lower than that of the EARS W2c method for all four cases, which indicates that

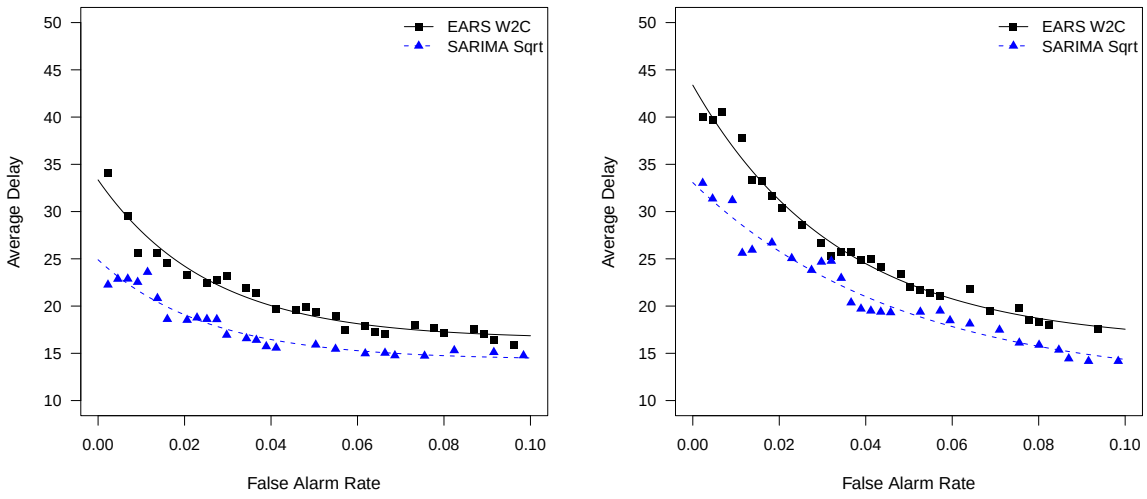
the SARIMA method detects outbreaks faster than EARS.

It is also noted that the differences between EARS and SARIMA in delays for BioSense_M1 (panel (C) and panel (D) in Figure3.1) are larger than those for AndNegEx_M2 (panel (C) and panel (D) in Figure 3.1). This difference mainly results from the fact that the forecast precision of the SARIMA method for high count series is better than for low count series [15]. Specifically, the average counts of BioSense_M1and AndNegEx_M2 in year 2002-2003 are 20.6, 4.9, respectively; after the square root transformation of the AndNegEx_M2 data series, the average is as low as 2.21, which limits the SARIMA forecast performance.

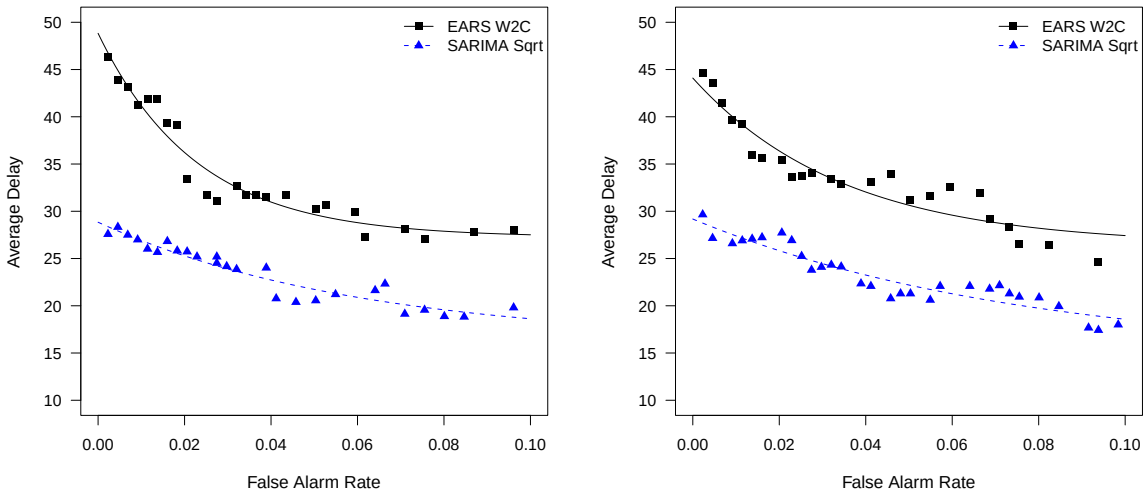
3.3 Analysis of CDC simulation data

To provide biosurveillance researchers a consistent and well-defined testbed for the objective comparison of aberration detection algorithms, the CDC publicizes 56 simulated datasets which are designed to simulate various natural sources [20]. Each of the 56 simulated background datasets is generated containing 1000 iterations of 6 years of daily data, 1994-1999, using a negative binomial distribution with superimposed outbreaks. These datasets include national and state pneumonia and influenza data and hospital influenza like illness. To match the known patterns of authentic syndromic data, adjustments are also made for days of the week, holidays, post-holiday periods, seasonality, and trend in these simulated background datasets [20, 21].

There are ten types of outbreaks which are randomly placed throughout the data series. Outbreak type one and two were 1-day spikes; outbreak types seven through nine and zero are based on the log normal distribution; and in outbreak types three through



(a) **(b)**



(c) **(d)**

Figure 3.1: AMOC Curves. Panel (a), (b), (c), (d) are the average delay versus false alarm rate for AndNegEx_M2 in 2002-2003,AndNegEx_M2 in 2003-2004, BioSense_M1 in 2002-2003 and BioSense_M1 in 2003-2004, respectively.

six, the lognormal distribution of attributable symptomatic cases is reversed[20].

In order to examine the surveillance performance of EARS W2c algorithm and SARIMA approach for a variety of data types and sizes, we also applied the methods on two sets of simulated background data and simulated outbreaks (s17_08 and s33_04) provided by the CDC. Here, the notation s17_08 indicates the 17th dataset of the 56 datasets with the 8th outbreak type. To make SARIMA approach work for CDC data series, both of these two data sets are selected because they show seasonality. Some characteristics of s17_08 and s33_04 are given in Table 3.2.

Table 3.2: Nature of CDC simulation datasets investigated

Data	Outbreak Type	Iteration	Trend	Seasonality	Mean	Standard Deviation
s17_08	Log Normal	732	Yes	Very	90.200	33.300
s33_04	Invert Log Normal	360	Yes	Medium	29.900	5.600

The background data sets generated by the CDC range from January 1, 1994 to December 31, 1999. We used the data in the year of 1997 to study the surveillance performance, so the data before the year of 1997 were treated as the baseline base to create the initial SARIMA model. The most suitable form of the SARIMA model fitted to each data series changed with the nature of the datasets and with the separation of weekdays and weekends. The SARIMA models for the four data series are given in Table 3.3.

Table 3.3: The SARIMA models for the CDC simulated datasets

Data	DOW	Seasonality	p	d	q	P	D	Q
s17_08	Weekday	260	0	1	1	1	0	1
s17_08	Weekend	104	1	0	1	0	1	1
s33_04	Weekday	260	1	0	1	1	0	1
s33_04	Weekend	104	1	0	1	0	1	1

The surveillance performance of the EARS and SARIMA methods are plotted in Figure 3.2. In terms of the surveillance metric timeliness, the SARIMA algorithm was su-

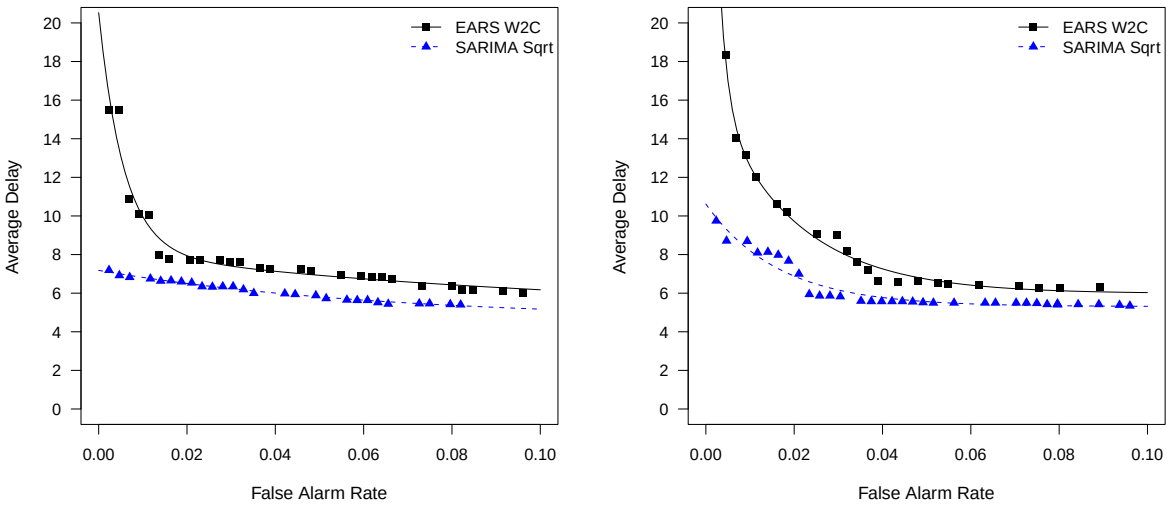


Figure 3.2: AMOC Curves. Panel (a), (b) are the average delay versus false alarm rate for s17_08 and s33_04 of CDC simulation data, respectively.

perior to EARS method across the practical levels of false alarm rates as shown in panel (A) and (B) in Figure 3.2. The average delay of the SARIMA method was much smaller than that of the EARS method at the lower false alarm rates. It was also noted that the average delay in detecting outbreaks in the CDC data was much smaller than that with the VAMHCS data. The primary reasons for this are (i) the latency period of the CDC outbreaks was 4, while the corresponding period for the VAHCS data was set at 6; (ii) more importantly, the injected case counts for the CDC outbreaks were large at the injection origin, while the injected counts of the VAHCS outbreaks were very low, as illustrated in Figure 2.1.

Chapter 4

Discussion

Modern biosurveillance research focuses on detecting disease outbreaks rapidly by developing surveillance systems which traditionally consist of three stages: data preparation, counts prediction, and aberration detection. The EARS W2c method itself incorporates two of these stages: count prediction using a 7-day moving average, and aberration detection by use of a standardized count and a threshold based on a desired recurrence interval value. The emphasis of our paper was to compare of the surveillance performance of the EARS W2c and SARIMA methods with respect to timeliness, sensitivity, and specificity. The use of the SARIMA method involves daily count prediction. To make the SARIMA modeling and forecasting be meaningful for this purpose, we selected data series with evident trend and seasonality from the CDC database of simulated background data. To compare EARS and SARIMA methods on the same basis, a series of thresholds were used to distinguish true positives from false alarms in the background data. The comparison between the EARS and SARIMA methods reflects their prediction performance for future counts.

If there exist fluctuations in the variance of count series, the logarithmic transformation is frequently applied. When developing SARIMA models for the data series with evidence of variance fluctuation, we tried to use logarithm-transformed data to cre-

ate SARIMA models. We attempted to model a data series with sparse counts (And-NegEx_M2_Temp), but the SARIMA technique could not be implemented for this case, which confirmed Williamson and Hudson's (1999) finding [15]. Furthermore, after investigating the nature of false alarms in background data series, we found that most of the false outbreak signals of EARS occurred in seasons with high variance, but this phenomenon seldom happened for the SARIMA method.

Our study suggests that without loss of sensitivity, well developed SARIMA model prediction can be used to detect outbreaks in a more timely way compared to the EARS W2c algorithm for the different types of case detection algorithms and time periods investigated. In the practical range of false alarm rates, as the background or baseline counts become higher this benefit of the SARIMA method becomes more apparent (see panel(C) and panel(D) of Figure 3.1). The results and findings based on the two data series provided by CDC were similar to those based on our datasets.

Algorithms for simulating outbreaks play an important role in evaluating health surveillance systems. Some statistical models, such as Poisson models and lognormal models are widely used by surveillance researchers. These types of models can describe the outbreak characteristics (such as the mean and variance) and are easy to study; however, they may not reflect all of the information in the baseline data. Our epidemic model for simulated outbreaks, based on authentic syndromic data in the Baltimore, Maryland area, involves complexity and realism in terms of factors such as the population, the age structure, the latency period, the severity level, the discovery rate, and so forth. Despite a loss of some information due to the aggregation of the data to a temporal data series, the simulated outbreaks based on the real daily counts and a carefully developed outbreak model allow us to make a meaningful comparison of different surveillance algorithms.

We successfully created the SARIMA model for our baseline data, and compared the

surveillance performance of the SARIMA and EARS algorithms. However, we noticed that there were some limitations and improvements needed in our research. First, we assumed that there were no known outbreaks in our data series. However, it is apparent that the data (e.g., in Figure 2.2) contain seasons of acute respiratory injections, such as influenza or respiratory syncytial virus, which affect the numbers of patient visits. This issue might affect the modeling and surveillance performance evaluation. Another improvement needed for our future research is to use a short guard band of two to four days for the updated forecasts in the SARIMA algorithm to avoid contamination by the gradual outbreak data. With this approach, some of the most recent data are not used to generate the forecasts. Second, the precision of SARIMA forecasts is sensitive to extreme values (i.e. outliers). To improve the reliability of the prediction, some outliers were replaced by the average of the past three daily counts. The criterion for identifying outliers needs to be better established. Finally, in the aberration detection stage, the statistics used for aberration detection are not independently and identically distributed. In our study, a signal that an outbreak had occurred was based on a single surveillance statistic value that corresponded to the count for the current day. Modifying the statistics by using cumulative sum (CUSUM) or exponentially weighted moving averages (EWMA) may lead to improvement in surveillance performance [22]. Since signals indicating that an outbreak has occurred may come in groups when there is no outbreak, other performance metrics may be useful in comparing the performance of competing methods [4, 23].

There is no signal method which can be universally applied to public health surveillance data to identify an unusually high number of cases of diseases or health events [15]. If several years of data and expertise are available, SARIMA model-based forecasts might enhance the performance of case count prediction in some cases, although more work and study are required.

Part II

GLR Control Charts for Poisson Data

Chapter 5

Introduction

Modern biosurveillance is the practice of monitoring a wide range of prediagnostic and diagnostic data to quickly detect, investigate and respond to disease outbreaks. In contrast to traditional biosurveillance which examines the existence of disease outbreaks retrospectively, modern biosurveillance uses less specific aggregated daily or hourly syndromic data to issue an emergency alarm in a timely manner [10]. These syndromic data are generally collected from prediagnostic health-care related sources, such as a count of the respiratory diagnoses from civilian office visits, the number of visits to military clinics, over-the-counter and pharmacy sales and school absence records. After filtering some underlying effects(e.g., seasonality, day-of-week) [24, 25], such count data series are often assumed to follow a certain discrete distribution, such as the Poisson distribution. If an undesirable event occurs, the rate of the Poisson distribution will change, which leads to a widespread use of statistical process control charts in public health monitoring, including Shewhart, CUSUM and EWMA charts [4, 24, 26, 27].

Statistical process control (SPC) involves using statistical methods to identify and analyze the variation in processes. In fact, every process displays variation. Some variation is natural and inherent to the process, which is referred to as common cause variation or non-assignable cause variation; other variation, however, may be caused by a fundamen-

tal change that is not supposed to be present in the process, which is known as special cause variation or assignable cause variation. Special causes can be eliminated from the process if they are identified and detected promptly. In practice, a process will be called in-control if there is no present special cause associated with this process. To identify whether a process is in-control or not, we often use a control chart, which is essentially a plot of a certain control statistic over time. If a control statistic falls outside limits, we will regard it as an out-of-control signal, which indicates some special causes may have been involved in the process, and we need to take actions to adjust the process back to normal. However, if a control chart gives a signal when the process is actually in-control, then the undesirable signal will be regarded as a false alarm. Thus, a control chart producing frequent false alarms has poor surveillance performance.

A number of control charts have been introduced. One of the popular control charts is the Shewhart chart, which was proposed by Walter A. Shewhart [28]. Since then, it has been widely used in a variety of areas due to its simplicity. It is known that the Shewhart chart has a good performance in detecting large process shifts, but it is rather insensitive to small shifts, because it only considers the most current data information. In contrast, the other two popular alternatives, cumulative sum (CUSUM) charts introduced by Page [29] and exponentially weighted moving average (EWMA) charts proposed by Roberts [30], can accumulate the past and current information for the process monitoring, which leads to an increase in the capability to detect small process shifts. However, CUSUM and EWMA charts generally are not as efficient in detecting large process shifts. For example, if the process mean is monitored, the EWMA chart with a small value of the smoothing parameter will give a good performance for detecting small shifts in the mean, while the CUSUM chart with a pre-determined tuning parameter close to the shift size will be good at detecting such small mean shifts. In practice, however, the shift size is usually unknown in advance. In this sense, control charts with the capability to detect a range of

shift sizes are more highly desired.

Several techniques have been developed to handle the issue of detecting a range of unknown shift sizes. One approach is to combine two or more control charts. For example, a Shewhart chart can be implemented in conjunction with a CUSUM chart, where the Shewhart chart is intended to detect large shifts while the CUSUM chart is designed to detect the small shifts quickly [31, 32]. Zhao *et al.* [33] proposed a method to combine two different CUSUM charts, called the dual CUSUM chart. Sparks [34] studied a multi-CUSUM chart approach, which employs multiple CUSUM charts simultaneously to detect a wide range of shifts, in which each chart is independently implemented with a different reference value. The control limit of each chart should be adjusted to give the desired in-control performance. A signal will be triggered if one of the multiple charts reaches its corresponding control limit. As the number of combined CUSUM charts increases, more control limits are then required, which leads to be an big disadvantage of applying this approach.

Another approach to effectively detect a wide range of shifts is to use adaptive CUSUM control charts, in which the process mean can be adaptively estimated with the data from the process, and the value of the CUSUM tuning parameter is then sequentially replaced by the mean estimate over time. Sparks [34] proposed an adaptive CUSUM (ACUSUM) chart using an EWMA statistic to estimate the process mean, based on which ACUSUM-C and WCUSUM-C charts are developed, see Jiang [35] and Shu [36] for details. On the one hand, adaptive chart performance could be enhanced by adjusting the multiple tuning parameters, but on the other hand, practitioners might find it difficult to select suitable parameter settings and determine an appropriate control limit.

The primary objective of this dissertation is to investigate another type of control chart with the ability to detect a wide range of shifts in the process mean. The proposed

chart is derived from a generalized likelihood ratio (GLR) statistic calculated from the Poisson distribution. The developments and properties of GLR charts for monitoring continuous variables have been proposed in literature such as Willsky *et al.* [37], Siegmund *et al.* [38], Lai [39, 40], Apley *et al.* [41], Capizzi [42], Hawkins *et al.* [43], Hawkins *et al.* [44], Han [45], Höhle [46], Reynolds *et al.* [47]. Although some researchers have applied the GLR algorithm to discrete data (e.g., Huang [48]), as a whole the GLR chart for attribute data has not been as well studied.

Under the research framework of SPC, the proposed GLR Poisson chart is intended for biosurveillance in which a popular surveillance approach developed by CDC, called EARS, has been widely used in U.S. and abroad currently [3, 12]. After developing the simulation system for this study, we have obtained some preliminary results, all of which indicate that compared with individual EARS, Shewhart, CUSUM and EWMA charts, the GLR chart has an overall good performance across a wide range of shifts in the Poisson mean. The superior surveillance performance of the GLR chart may result from the fact that the GLR statistic is maximized over likelihoods of all the possible shift sizes.

The specific goals of the GLR Poisson chart research can be summarized as follows:

1. Develop the GLR Poisson chart for various magnitudes of the in-control mean, and investigate the performance consistency of the GLR chart when varying the in-control mean;
2. Compare the performance between the GLR chart and other individual charts;
3. Demonstrate how to create a GLR chart and how to interpret the results;
4. Provide guidelines on designing GLR charts in practice;
5. Provide reference tables of control limits for varying in-control performance;

6. Develop a R package to facilitate use of the GLR chart, and publish it in the Comprehensive R Archive Network (CRAN).

In Chapter 6, we introduce some background involved this study and some traditional control charts. In Chapter 7, we derive the statistics of the GLR chart and discuss the properties of the statistics. In Chapter 8, we illustrate how to choose the window length of the GLR chart and how to determine the control limits. We also provide examples of calculations and plots for the GLR chart. In Chapter 9, we compare the performance between the GLR chart and other alternative charts under a variety of circumstances. Last of all, in Chapter 10 we give our conclusions regarding the GLR chart and propose future research directions.

Chapter 6

Background description and control charts

6.1 Background description

To timely detect and promptly respond to potential disease outbreaks, modern biosurveillance systems widely incorporate syndromic data sources including numbers of emergency room visits, school absentee rates, numbers of visits to military clinics. Such data series are often assumed to follow a certain discrete distribution, for example, the Poisson distribution. Consequently, statistical approaches and SPC charts are widely incorporated in biosurveillance. However, due to underlying characteristics of the syndromic sources, such as seasonality and day-of-week effect, and uncontrollable variations involved in the data, the implementation of statistical methods for biosurveillance is not as straightforward as for the industrial process monitoring, so surveillance methods with flexibility for the data sources and simplicity for the public health practitioners are desired.

As mentioned in Part I, in general, health-related surveillance systems involve the following stages: data preparation, count prediction, aberrancy detection, and outbreak signal evaluation. SPC charts are often involved in the stage of aberrancy detection. Comprehensive reviews of the applications of SPC methods to health surveillance can

be found in Woodall [4] and Tsui *et al.* [49]. Although this dissertation research of surveillance methods is ultimately for health care surveillance or similar areas, the method development and evaluation will be primarily under the SPC research framework.

For the process monitoring situations discussed in this part, our primary surveillance characteristic is the mean of Poisson count, which is denoted by λ. The goal is to effectively detect any increase in λ from its in-control mean λ_0 to an out-of-control value λ_1.

Although the collected syndromic data are sometimes correlated over time, for simplicity, we assume that they are independent variables in this study. Suppose we monitor a sequence of independent Poisson random variables over time $\{X_1, X_2, \cdots, X_k\}$, in which the true mean of X_i is λ_i, and in the normal conditions, we have $\lambda_i = \lambda_0$, but after an undesired event occurs at an unknown time point τ^* between sample τ and $\tau + 1$, where $\tau < k$ and assume the time interval is 1, the values of λ_i change from λ_0 to λ_1. The corresponding hypotheses test can be expressed as

$$\begin{aligned} H_0: &\quad \lambda_i = \lambda_0 \qquad \text{for all } i \geq 1 \\ H_1: &\quad \lambda_i = \begin{cases} \lambda_0 & \text{if } 1 \leq i \leq \tau \\ \lambda_1 & \text{if } i > \tau \end{cases} \qquad \text{where } \lambda_1 \neq \lambda_0. \end{aligned} \tag{6.1}$$

For this case, changes in λ may be an increase or a decrease, so a control chart based on these hypotheses is a two-sided control chart. In some applications, for example, the biosurveillance for the mortality and morbidity rate, the interest is only to detect an increase in the rate, so we change the restriction for λ_1 in Equation (6.1) from $\lambda_1 \neq \lambda_0$ to $\lambda_1 > \lambda_0$, based on which the derived control chart is a one-sided control chart.

In general, the shift of the process mean appears in some pattern, such as sustained

shifts, transient shifts and drifts. In some specific applications, the shift may follow a particular distribution. For example, in biosurveillance, the gamma distribution and the lognormal distribution are often used to describe the shift pattern of some infectious diseases [50, 51, 52, 53]. In the current study, we suppose that the pattern is a persistent jump change (i.e.,sustained shift) in the mean.

When the process mean changes from the in-control condition to the out-of-control condition, in general, there are two types of measures of the shift size, i.e., the proportional shift and standardized shift. Specifically, if μ_0 and σ_0 are the mean and standard deviation under the in-control condition, and μ is the process mean under the out-of-control condition, then the proportional shift δ_p and standardized shift δ_s can be represented as

$$\delta_p = \frac{\mu - \mu_0}{\mu_0}, \quad \text{and} \quad \delta_s = \frac{\mu - \mu_0}{\sigma_0}, \tag{6.2}$$

respectively. For the Poisson distributed data, the variance and mean are equal, thus, the relationship between these two shifts is

$$\delta_s = \sqrt{\lambda_0}\delta_p. \tag{6.3}$$

One type of shift for Poisson data, consequently, can be easily changed to another type of shift. If the data are normally distributed, in general, the standardized shift is employed in the statistical control literature [35, 44, 47, 54]. For the Poisson distributed data, however, both the proportional shift and the standardized shift have appeared in the literature [55, 56, 57]. In this study, we use the standardized shift, which makes it possible to easily see if the GLR chart performance for Poisson data is consistent to that for normal data [47].

In general, the performance of a detection method can be evaluated based on two

primary measurements: the average time to signal when the process is in-control (ATS_0), and the conditional expected delay (CED) when the process is out-of-control. Different detection methods are compared by their values of CED at the same value of ATS_0. The ATS_0 is the expected time from the start of monitoring to the time point where a signal is triggered, so the ATS_0 can be viewed as a measure of the false alarm rate (FAR). Let T be the time that a signal triggers and τ be the time of the onset of the mean change. If the process is in-control, we want T to be as large as possible so that the false alarm rate is minimized. We can adjust the control limit to approach this specified ATS_0 even without any change in the process. In general terms, an algorithm with a smaller ATS_0 has higher sensitivity for detecting anomalies, though this comes at the expense of an increased false alarm rate.

On the other hand, the detection delay is defined as the time interval from the mean shift to the signal. We assume that the shift time τ^* is uniformly distributed in the interval $[\tau, \tau + 1]$. If T is the time point where a signal triggers, then the delay to detect the shift is $T - \tau^*$. The monitoring goal is usually to minimize the expected detection delay over a range of process shifts while requiring a specified value of ATS_0. The measured detection delay can either be a zero-state or a steady-state detection delay. The zero-state detection delay is based on sustained shifts in the monitoring parameter that occur under the initial startup conditions of the control charts, while the steady-state delay is based on delayed shifts in the parameter. Since the stopping time T is a random variable, we use the conditional expected delay (CED) defined as $E(T - \tau^* | T \geq \tau^*)$ to evaluate the detection performance of different methods. The CED metric is the better measurement of performance for the comparisons of the various methods considered in this study. The reason for using CED involves the fact that the EWMA chart has a slight advantage over the CUSUM chart at early stages of monitoring. At the initial stages of monitoring, the variance of the EWMA statistics is smaller, which leads to tighter control limits. As a re-

sult, the one-sided EWMA has a slight head start over the competing CUSUM charts. In general, the steady-state assumptions seem more realistic.

Poisson data are count numbers, so the distribution's discrete feature has some effects on the desired ATS_0. For example, if a control statistic is only based on the current observation and the desired $\text{ATS}_0 = 1500$ when $\lambda_0 = 2$, theocratically, this value of ATS_0 cannot be obtained for an upward-sided chart. An ATS_0 value of 1500 can be viewed as one false alarm every 1500 observations on average when the process is in-control, which means the corresponding false alarm rate is around $1/1500$. Also, the probability of an extreme observation (say $> x$) occurs can be measured by $1-\text{CDF}$, where CDF is the cumulative density function valued at x, i.e., $\text{CDF} = \Pr(X \leq x)$. Thus, if a signal is triggered by an observation x_k, where $x_k > x$, then the relationship among CDF, FAR, and ATS_0 is approximately represented as [58]

$$\text{ATS}_0 \approx \frac{1}{\text{FAR}} \approx \frac{1}{1-\text{CDF}}. \tag{6.4}$$

For the Shewhart method, the relationship shown in Equation (6.4) is equality instead of approximation. Table 6.1 gives the values of CDF, FAR and ATS_0 for Poisson data with $\lambda_0 = 2$. We can see that ATS_0 jumps from 911.81 to 4211.46 when the control limit changes from 7 to 8, so we cannot have the desired $\text{ATS}_0 = 1500$ for this case.

Table 6.1: CDF for the Poisson distribution with $\lambda_0 = 2$

x	CDF	FAR	ATS_0
0	0.13534	0.864665	1.16
1	0.40601	0.593994	1.68
2	0.67668	0.323324	3.09
3	0.85712	0.142877	7.00
4	0.94735	0.052653	18.99
5	0.98344	0.016564	60.37
6	0.99547	0.004534	220.57
7	0.99890	0.001097	911.81
8	0.99976	0.000237	4211.46
9	0.99995	0.000046	21506.27
10	0.99999	0.000008	120362.66
11	1.00000	0.000001	732807.34

6.2 Control charts

6.2.1 EARS W2c approach

As mentioned in Part I of this dissertation, the CDC BioSense program developed the early aberration reporting system (EARS), which is designed to detect increases in counts or rates with health-related data. EARS method has been widely used in various cities, counties, and state public health officials in the United States and abroad [3, 12].

The EARS consists of several methods, including EARS W2c. The observed case counts, e.g., the numbers of visits at a particular hospital emergency department with a specific syndrome, are separated into weekdays and weekends/holidays first, and then the counts of weekdays and weekends/holidays are treated as different time series data with separate day indices. The case of weekdays is taken as an example to illustrate how the W2c statistics are calculated. Let k be the time series index, and x_k be the observed case counts on day k. Note that there is no assumption on the distribution of x_k. The W2c

statistic is expressed as

$$W_k = \frac{x_k - \overline{x}_k}{S_k}, \tag{6.5}$$

where $\overline{x}_k$ and S_k are the moving sample mean and standard deviation calculated with a 2-day lag respectively, i.e.,

$$\overline{x}_k = \frac{1}{7}\sum_{i=k-3}^{k-9} x_i, \qquad S_k = \sqrt{\frac{1}{6}\sum_{i=k-3}^{k-9}[x_i - \overline{x}_k]^2}. \tag{6.6}$$

The value of is replaced by 1 if $S_k < 1$.

The W2c method signals if the statistic in Equation 6.5 exceeds a given threshold. The thresholds such that the resulting value of the recurrence interval (RI) is a specified value can be found in the reference table provided by CDC BioSense [13]. Although the EARS method seldom appears in SPC literature, for the purpose of biosurveillance, we compare its performance with other alternative approaches considered in this study, and the threshold of EARS is determined using simulations to obtain the specified ATS_0 instead of using those thresholds given by CDC.

6.2.2 Shewhart control charts

In contrast to the EARS which is popular in biosurveillance, the Shewhart control chart, originally proposed by Shewhart [28], is most widely used in industrial SPC. The Shewhart $\overline{x}$ chart for monitoring the mean μ is based on plotting the sample mean $\overline{x}_k$ versus time k, and a signal is given at time k if $\overline{x}_k$ falls outside of its control limits $\mu_0 \pm h_S\sigma_0/\sqrt{n}$, where μ_0 is the in-control mean, σ_0 is the standard deviation and h_S is the control limit. In practice, the control limit h_S is usually taken to be 3 to yield the traditional "three-sigma" control limits. In our study, the control limit is set to give the specified in-control perfor-

mance. More discussion on the Shewhart chart can be found in the literature [59, 60, 61]. For the Poisson data, the monitoring statistic of the Shewhart chart for the sample k is a count number, so the control chart statistic can be written as

$$S_k = \frac{x_k - \lambda_0}{\sqrt{\lambda_0}}. \tag{6.7}$$

As stated by Woodall [62], Poisson data with a small mean are skewed to the right, the traditional k sigma limits may be inappropriate. Shewhart control charts are useful for application in phase I [61], though they are relatively insensitive to small process shifts occurring during phase II. To improve the performance of the Shewhart chart, the monitoring statistic can be modified to accumulate the information across past samples. Champ *et al.* [60] evaluated the runs rules method, which is based on patterns of points in a Shewhart chart, and using runs for improving the ability to detect small shifts. However, a better approach may be to use a CUSUM or EWMA control chart [60].

6.2.3 CUSUM control charts

The CUSUM chart was proposed by Page [29], and since then its properties and applications have been studied and discussed extensively in literature [44, 51, 62, 63, 64, 65]. In the context of biosurveillance, the CUSUM chart was applied by Hill *et al.* [66] and Weatherall and Haskey [67] primarily in surveillance of congenital malformations, and was recommended by Hutwagner [68] in surveillance for Salmonella outbreaks. Recently, Rogerson [24] used CUSUM charts for spatial surveillance when incidence rates varied over time.

The CUSUM chart for monitoring the process mean is equivalent to applying a se-

quence of sequential probability-ratio tests (SPRTs) [29, 65] for testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu = \mu_1$, 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, even if the actual shift is unknown. In contrast, μ_1 in the GLR charts is sequentially estimated from the process data, as will be discussed later.

In contrast to the Shewhart chart which only uses the information given by the current sample, the CUSUM chart can incorporate information given by past samples using the cumulative sums of the log likelihood ratio statistics. The Poisson CUSUM chart can be formulated from a sequence of SPRTs for testing $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda = \lambda_1$. Let r_k be the increment of log likelihood ratio that is accumulated for the CUSUM Poisson statistic at sample k, then r_k can be calculated by

$$r_k = \ln\frac{f(x_k|\lambda_1)}{f(x_k|\lambda_0)} = x_k \ln\frac{\lambda_1}{\lambda_0} - (\lambda_1 - \lambda_0), \tag{6.8}$$

where λ_1 is the tuning parameter that specifies the particular shift size for which the Poisson CUSUM chart is designed to respond quickly. The CUSUM chart will produce a signal if the difference between the cumulative increment of the current sample and that of previous minimum exceeds a pre-determined control limit. Let C_k be the difference of cumulative increments and h_C be the control limit of the Poisson CUSUM chart, then the Poisson CUSUM chart signals if $C_k > h_C$, where

$$C_k = \sum_{i=1}^{k} r_k - \min_{n \le k} \sum_{j=1}^{n} r_j. \tag{6.9}$$

The CUSUM chart statistic is usually represented in an iterative form

$$C_k = \max\{0, C_{k-1}\} + r_k = \max\{0, C_{k-1}\} + x_k \ln\frac{\lambda_1}{\lambda_0} - (\lambda_1 - \lambda_0), \tag{6.10}$$

where $C_0 = 0$ and $k = 1, 2, \cdots$, which means it can be simply expressed in the form of the previous accumulation C_{k-1} and the current increment r_k. In general, the control statistic for monitoring the mean is expressed in a different form by dividing both sides of Equation (6.10) by the constant $\ln(\lambda_1/\lambda_0)$. If $\lambda_1 > \lambda_0$, then leads to a modified CUSUM statistic C'_k, given by

$$C'_k = \max\{0, C'_{k-1}\} + x_k - \frac{\lambda_1 - \lambda_0}{\ln\frac{\lambda_1}{\lambda_0}}, \tag{6.11}$$

where $(\lambda_1 - \lambda_0)/\ln(\lambda_1/\lambda_0)$ is usually referred to as the reference value of the CUSUM statistic. In this case, a signal will be given if $C'_k > h'_C$, where $h'_C = h_C/\ln(\lambda_1/\lambda_0)$. Since the increment associated with the current sample in Equation (6.11) is straightforward, C'_k values are frequently used in practice [32, 36, 55, 69]. In this study, our interest is the upward shift in the Poisson mean, so one-sided CUSUM charts are used. We apply C'_k in our simulations. For example, the control limits of CUSUM charts displayed in tables are in the form of h'_C, whereas some discussion about the relationship between GLR charts and CUSUM charts are based on C_k due to their similar expression.

The CUSUM chart discussed above is a one-sided chart for monitoring a mean shift from $\lambda = \lambda_0$ to a specified λ_1. In practice, a two-sided CUSUM chart can be obtained by using two one-sided CUSUM charts together in combination, but control limits won't usually be the same for a skewed distribution.

6.2.4 EWMA control charts

In addition to the CUSUM chart, another control chart that can incorporate the past information is the EWMA chart, which was originally proposed by Roberts [30]. Since then, many authors have evaluated its properties and performances, including Crowder [70, 71], Lucas *et al.* [72], Yashchin [73] and Borror *et al.* [74].

The EWMA control chart is plot of a weighted average of the current and past samples results over time, and is efficient in detecting small to moderate shifts in the process parameter. The monitoring statistic of the EWMA chart at sample time k can be expressed as

$$E_k = \alpha x_k + (1 - \alpha) E_{k-1}, \tag{6.12}$$

where α is the smoothing parameter with a range of $0 < \alpha \leq 1$. In general, one uses small values of α to detect small parameter shifts, and use a large α to detect large shifts. If $\alpha = 1$, then the EWMA chart is equivalent to the Shewhart chart, since only the current observation is considered. In general, the starting value E_0 is set at its target value, and values of α between 0.05 and 0.3 are often applied in practice.

If the EWMA chart is intended to detect the upward shift in the mean, then a modified surveillance statistic was recommended by Crowder *et al.* [75] to prevent the EWMA statistic from taking the very small values which may lead to serious inertial problems, especially if the value of α is small [76]. The modified approach is to use a reflecting barrier at $E_k = \lambda_0$, and then the upward-sided monitoring statistics can be expressed as

$$E_k^+ = \max\{E_k, \lambda_0\}. \tag{6.13}$$

In our study, we use E_k^+ as the EWMA statistic and the control limit is determined to give the specified in-control performance using simulations.

In general, the properties of the EWMA chart are similar to that of the CUSUM chart in terms of the ability to detect shift in the parameters being monitored. Recently, Han *et al.* [55] evaluated the performance of CUSUM and EWMA charts for biosurveillance, and concluded that the CUSUM chart is better than the EWMA chart when detecting a large shift with a later changepoint, while the EWMA chart outperforms the CUSUM chart in

situations with small shift and an early changepoint. However, there was no reflecting barrier for their one-sided EWMA charts, which makes their setting different from ours. Fricker [77] studied the Shewhart, CUSUM and EWMA charts for monitoring the shift in the mean of normal data in the biosurveillance context. The author stated that the EWMA chart had the shortest CED values for any shift size, followed by the CUSUM chart, and then the Shewhart chart.

6.2.5 GLR control charts

The use of the generalized likelihood ratio test (GLRT) in the process change point detection was first proposed by Lorden [78]. Incorporating the GLRT into the CUSUM working mechanism, the GLR approach is to doubly maximize the process parameter and all possible change points at each sampling time. More specifically, the GLR statistic uses an estimated size of shift that is evaluated at each possible time of change and gives the maximum likelihood. Then based on this estimated shift size a modified likelihood ratio test is conducted, which is developed for testing H_0, that the process parameter remains unchanged, against H_1, that parameter has shifted to some unknown out-of-control value estimated by the maximum likelihood estimator.

The GLR algorithm can be described statistically. Suppose the Poisson mean has been changed from the in-control mean λ_0 to an unknown mean λ_1 since a certain time point, and $\beta_j^k(\lambda_1)$ is the log likelihood ratio for the samples from time j up to time k, then $\beta_j^k(\lambda_1)$ can be given by

$$\beta_j^k(\lambda_1) = \sum_{i=j}^{k} \ln \frac{f(x_i|\lambda_1)}{f(x_i|\lambda_0)}. \tag{6.14}$$

Therefore, this ratio is a function of two unknown parameters: the change time and the mean after the change. The standard statistical approach is to use the maximum likeli-

hood estimates of these two parameters. That is

$$R_k = \max_{1 \le j \le k} \sup_{\lambda_1} \beta_j^k(\lambda_1), \tag{6.15}$$

where the R_k is the surveillance statistic of the GLR algorithm. The techniques used here are often regarded as a *double* maximization.

The derivation of the GLR statistic leads to two inherent advantages over other control charts: (i). the built-in change point estimator that provides the important information about when a change has actually occurred in the process; (ii). double maximization of log-likelihood ratio over all possible change points at each sampling time, that makes it possible to detect a wide range of shift sizes. On the other hand, double maximization leads to more intensive computational efforts to generate the GLR statistic. As pointed out by Basseville *et al.* [79], without a recursive formulation of the GLR statistic, its practical implementation was not always possible. Fortunately, much work has been done to reduce the computational burden involving the GLR algorithm; see Basseville *et al.* [79] for a comprehensive review of the simplification and modification for this changepoint detection algorithm.

A natural modification to get around the computational difficulty is to use a moving window with size of m, over which the maximization of the log-likelihood ratio is only taken within the past m samples. Such a window-limited GLR scheme was first introduced by Willsky *et al.* [37]. Since then, this approach has been popularly used. Basseville *et al.* [80] proposed another modification of the GLR algorithm, which involved mixing the GLR detector given by Willsky *et al.* [37] and CUSUM test given by Hinkley [81], and stated that this algorithm is far less time consuming, while still efficient and robust.

Siegmund *et al.* [38] studied detecting shifts in in the normal mean with known

variance as a special case in sequential change point detection using the GLR algorithm. They derived the asymptotic average run length as the control limit goes to infinity, displayed some numerical results using Monte Carlo simulations, and stated that the GLR chart is superior to a standard CUSUM chart for detecting a wide range of mean shifts. They found there are only minor differences between the GLR chart and the combined Shewhart-CUSUM charts. They also considered situations such as monitoring the mean change when the variance is unknown, as well as when the data are multivariate or autocorrelated.

In the context of manufacturing, Apley *et al.* [41] compared the performance of GLR, Shewhart and CUSUM charts for monitoring the uncorrelated residuals of the appropriated ARIMA model. In the case of sustained mean shifts in the original process, they found the performance of these three charts varied with the settings. The first setting assumed that fault signatures had small steady-state magnitude but obvious transient dynamics. Here the GLR chart outperformed other two charts. The second setting assumed fault signatures had a relatively large steady-state magnitude and insignificant transient dynamics. Then the CUSUM chart was better than the GLR chart, and the Shewhart chart had very poor performance. The last setting was for a fault signature with neither a large steady-state magnitude, nor significant transient dynamics. They found that none of the three charts performed very well. Also for monitoring fault signatures, Runger *et al.* [82] evaluated the performance of the Cuscore chart and the GLR chart using the sine wave and the linear trend fault signatures, and they stated that the GLR charts had a better, or at least the same, performance than the Cuscore chart.

More recently, Reynolds *et al.* [47] evaluated the performance of the GLR chart for monitoring a normal mean process with three types of shifts: sustained shifts, transient shifts and drifts in the mean. The window-limited GLR scheme was used to reduce the

computational burden involved in the GLR statistic, and a window length providing a nearly optimal performance was recommended for practice. Detailed guidance for calculating the GLR statistic and for plotting the GLR chart with some illustrative examples were provided as well. In addition to deriving the theoretical relationship between the CUSUM chart and the GLR chart, they also made numerical comparisons of the GLR chart with alternatives, such as the Shewhart chart, CUSUM chart, combinations of traditional charts and adaptive charts. They found that across a wide range of mean shifts, the overall performance of the GLR chart is at least as good as the alternatives, with a benefit that designing a GLR chart only requires the moving window size and the control limit.

In contrast to the straightforward approach to obtain the control limit by using a linear relationship in Reynolds *et al.* [47] paper for the normal mean, Lai [39] pointed out that it was very difficult to find a satisfactory choice of both window size and control limit for detecting additive changes in a state-space model. Considering this difficulty, Capizzi [42] proposed a completely automatic design procedure to estimate the control limit that satisfied a constraint on the mean time between false alarms.

Motivated by the work done by Siegmund *et al.* [38] and Apley *et al.* [41], Han *et al.* [45] modified the stopping time rule of the GLR algorithm, so their GLR statistic is slightly different from that one given by Apley *et al.* [41]. In some sense, while the model used by Apley *et al.* [41] emphasizes the detection of the beginning change pattern, Han *et al.* [45] concerns the end change pattern. In fact, the GLR statistic given by Siegmund *et al.* [38] is a special case of Han *et al.*'s [45] version. In that paper, they theoretically proved that for most cases the GLR chart had the best performance among three compared charts, i.e., Cuscore, GLR and CUSUM, in detecting any mean change, and also presented some numerical work. Note that there is some difference between the results given by Han *et al.* [45] and Reynolds *et al.* [47] due to the difference in their methods.

In addition to the focus on the GLR chart itself, Kay *et al.* [83] established an invariance property of the generalized likelihood ratio test (GLRT) in light of its widespread use. They pointed out that the GLRT is invariant with respect to transformations for which the hypothesis testing problem itself is invariant. The theoretical proof for the property is followed by some realistic examples in their paper.

Although the GLR algorithm has become popular, the current methods have been focused mostly on monitoring continuous data, especially normally distributed data. The GLR chart for monitoring the count data, such as negative binomial data and Poisson data, have not been thoroughly studied. In biosurveillance, the GLR chart for monitoring Poisson data is particularly important not only because some popular SPC methods need pre-specified parameters, but because the GLR method has an ability to detect a wide range of mean shifts.

6.3 Metrics of performance evaluation

For a specific shift in the mean, the chart performance can be measured by its value of CED at a specified value of the ATS_0. However, the chart performance often depends on the shift size of the process mean. For example, EWMA charts perform better in detecting small mean shifts than in detecting large mean shifts. In contrast, the Shewhart chart is more efficient to detect large mean shifts than small mean shifts. The size of mean shifts is usually unknown in advance, so a fair amount of work has been devoted recently to developing methods for evaluating control charts with unknown shifts [33, 47, 84, 85], for which an index for evaluating the overall performance over a range of shifts is needed. As Reynolds *et al.* stated [47], an ideal index should include the following effects:

1. The loss per unit time at a particular shift;

2. The delay of detecting this shift;

3. The likelihood associated with this shift.

One of such indices, called extra quadratic loss (EQL), has been used in a number of papers [32, 47, 85, 86, 87]. More specifically, let x_t be the Poisson process observation at the time point t between the shift time point τ^* and the signal time point T, then the quadratic loss for Poisson data can be formulated as

$$\begin{aligned} E\Big[\Big(\frac{x_t-\mu_0}{\sigma_0}\Big)^2\Big] &= \frac{E(x_t-\mu_0)^2}{\sigma_0^2} + \frac{(\mu_1-\mu_0)^2}{\sigma_0^2} = \frac{\lambda_1}{\lambda_0} + \delta^2 \\ &= \delta^2 + \frac{1}{\sqrt{\lambda_0}}\delta + 1, \end{aligned} \tag{6.16}$$

so the extra quadratic loss (EQL) is $\delta^2 + \delta/\sqrt{\lambda_0} + 1 - 1 = \delta^2 + \delta/\sqrt{\lambda_0}$, and the extra quadratic loss function for a specified mean shift δ is then calculated as

$$EQL(\delta) = (\delta^2 + \delta/\sqrt{\lambda_0}) \cdot E(T-\tau^*) \tag{6.17}$$

$$= (\delta^2 + \delta/\sqrt{\lambda_0}) \cdot CED. \tag{6.18}$$

If the mean shift δ follows a specified prior distribution $\pi(\delta)$, then the EQL for Poisson data over the range $\delta \in [a, b]$ can be represented as

$$EQL = \int_a^b EQL(\delta)\pi(\delta)\mathrm{d}\delta = \int_a^b (\delta^2 + \delta/\sqrt{\lambda_0}) \cdot \pi(\delta) \cdot CED\mathrm{d}\delta. \tag{6.19}$$

Since the EQL considers the shift pattern, the loss for a particular shift, and the detection delay, the EQL can be used to measure the overall chart performance when the mean shift varies.

Other than the EQL, some authors [33, 54, 85] proposed alternative approaches to measuring the overall chart performance when shift sizes are unknown, one of which is called expected relative average run length (ERARL), which measures the relative performance between two methods in terms of the out-of-control performance. Let $ARL_1(\delta)$ and $ARL_2(\delta)$ be out-of-control average run lengths (ARL) of two charts, respectively, and the shift range be $[a, b]$, then the ERARL is defined as

$$ERARL = \int_a^b \frac{ARL_1(\delta)}{ARL_2(\delta)} \cdot \pi(\delta) \mathrm{d}\delta. \tag{6.20}$$

To let the notation be consistent in this study, we change the term ERARL to ERCED standing for the expected relative CED. Then the ERCED can be expressed as

$$ERCED = \int_a^b \frac{CED_1(\delta)}{CED_2(\delta)} \cdot \pi(\delta) \mathrm{d}\delta. \tag{6.21}$$

The value of ERCED can be viewed as a performance index of two control charts. Specifically, the larger value of ERCED, the worse the relative performance is for the first control chart.

In addition, we use two types of the prior distributions of δ, i.e., a gamma distribution and an uniform distribution. As mentioned earlier, the gamma distribution is often used to describe a shift size in biosurveillance. There are two parameters associated with the gamma distribution, i.e., the shape parameter and the scale parameter. We use the notation $G(a, b)$ to represent a gamma distribution with shape parameter a and scale parameter b.

For the gamma distribution, the value of δ is continuous in a range, say $\delta \in [a, b]$ over which the gamma density is positive. We evaluate the integral using the numerical

quadrature rule. For example, EQL in Equation (6.19) can be rewritten as

$$EQL = \int_a^b EQL(\delta)\pi(\delta)\mathrm{d}\delta = \sum_{i=1}^{n} w_i EQL(\delta_i)\pi(\delta_i), \tag{6.22}$$

where n is the total number of δ values used, w_i is the weight assigned to the corresponding δ_i. In practice, there are many choices of w_i, one of which is based on the Gaussian quadrature rule. In numerical analysis, a quadrature rule is an approximation of the definite integral of a function, usually stated as a weighted sum of function values at specified points within the domain of integration. More details about the Gaussian quadrature rule can be found in book *Introduction to Numerical Analysis* [88].

In the current study, we used 11 levels of sustained shift sizes in total to represent a wide range of shifts, which can be found in the second column in Table 6.2. The column titled original nodes and original weights are the values of Gaussian nodes and weights in the range $[-1, 1]$, respectively. When applying Equations (6.19) and (6.21), we conducted additional simulations to obtain the values of CED, in which shift sizes are the transformed nodes listed in column titled transformed weights, and then the corresponding transformed weights in the last column are used for the summation in Equation (6.22).

To evaluate chart performances in a variety of settings, besides varying the shift sizes, we also changed the values of ATS_0 with a range from 100 up to 1500. In biosurveillance, the activity monitoring operating characteristic (AMOC) curve is a plot of the conditional expected delay (CED) as a function of $1/ATS_0$, and is an useful and popular method for assessing the performance of algorithms that detect outbreaks of disease when a range of ATS_0 is considered[89]. More discussion on AMOC curves can be found in the biosurveillance literature [90, 91, 92, 93].

Table 6.2: Nodes and weights based on the Gaussian quadrature rule

	Shift size	Original nodes	Original weights	Transformed nodes	Transformed weights
1	0.25	-0.978	0.056	0.323	0.188
2	0.50	-0.887	0.126	0.631	0.424
3	0.75	-0.730	0.186	1.161	0.629
4	1.00	-0.519	0.233	1.873	0.787
5	1.50	-0.270	0.263	2.715	0.887
6	2.00	0.000	0.273	3.625	0.921
7	3.00	0.270	0.263	4.535	0.887
8	4.00	0.519	0.233	5.377	0.787
9	5.00	0.730	0.186	6.089	0.629
10	6.00	0.887	0.126	6.619	0.424
11	7.00	0.978	0.056	6.927	0.188

Chapter 7

Derivation of GLR control charts

Similar to the traditional SPC control charts, a GLR control chart is a plot of the GLR chart statistics against the sample time. In this chapter, we first derive the GLR chart statistic for the Poisson data. We then explore the features and properties of the derived statistics using graphical, mathematical and statistical approaches. In the following section, we modify the two-sided GLR statistics to one-sided statistics with a mathematical function. We give discussion and some suggestions about the use of GLR charts throughout the chapter.

7.1 Derivation of GLR chart statistics

A SPC chart is generally a plot of the control statistic as a function of the sample time. If the statistic falls outside a specified threshold or control limit, an out-of-control signal is triggered. In this section, we derive the GLR chart statistics based on the generalized likelihood ratio of the Poisson data.

Let X be the Poisson random variable and λ be the mean, then its mass function is $P(x;\lambda) = \lambda^x e^{-\lambda}/x!$, and its variance is λ. If a sequential sample $\{X_1, X_2, \cdots, X_k\}$ is

obtained, and we assume that mean shift occurs from λ_0 to an unknown mean λ_1 at time point τ^*, where $\tau < \tau^* < \tau + 1 \leq k$, then the likelihood function at sample k can be calculated as

$$L(\tau, \lambda_1 | x_1, x_2, \cdots, x_k) = \frac{\lambda_0^{\sum_{i=1}^{\tau} x_i} \lambda_1^{\sum_{i=\tau+1}^{k} x_i} e^{-\tau\lambda_0} e^{-(k-\tau)\lambda_1}}{\prod_{i=1}^{k} x_i!}. \tag{7.1}$$

On the other hand, if the mean stays in-control until the sample k, then the likelihood function at sample k can be expressed as

$$L(\infty, \lambda_0 | x_1, x_2, \cdots, x_k) = \frac{\lambda_0^{\sum_{i=1}^{k} x_i} e^{-k\lambda_0}}{\prod_{i=1}^{k} x_i!}. \tag{7.2}$$

If there has been a shift in the mean to some unknown λ_1 between τ and $\tau + 1$, then the shifted mean can be estimated by the maximum likelihood estimator, given by

$$\hat{\lambda}_{1,\tau,k} = \frac{\sum_{i=\tau+1}^{k} x_i}{k - \tau}. \tag{7.3}$$

To determine whether there is a mean shift, the GLR Poisson control statistic is then derived from the following log likelihood ratio:

$$\begin{aligned}
R_k &= \ln \frac{\max\limits_{0 \leq \tau < k, 0 \leq \lambda_1 < \infty} L(\tau, \lambda_1 | x_1, x_2, \cdots, x_k)}{L(\infty, \lambda_0 | x_1, x_2, \cdots, x_k)} \\
&= \ln \frac{\max\limits_{0 \leq \tau < k} \left[\hat{\lambda}_{1,\tau,k}^{\sum_{i=\tau+1}^{k} x_i} e^{-(k-\tau)\hat{\lambda}_{1,\tau,k}} \right]}{\lambda_0^{\sum_{i=\tau+1}^{k} x_i} e^{-(k-\tau)\lambda_0}} \\
&= \ln \Big[\max_{0 \leq \tau < k} \Big[\Big(\frac{\hat{\lambda}_{1,\tau,k}}{\lambda_0} \Big)^{\sum_{i=\tau+1}^{k} x_i} e^{-(k-\tau)(\hat{\lambda}_{1,\tau,k} - \lambda_0)} \Big] \Big] \\
&= \max_{0 \leq \tau < k} \Big[\ln \Big(\frac{\hat{\lambda}_{1,\tau,k}}{\lambda_0} \Big) \sum_{i=\tau+1}^{k} x_i - (k - \tau)(\hat{\lambda}_{1,\tau,k} - \lambda_0) \Big].
\end{aligned} \tag{7.4}$$

Using Equation (7.3), R_k in Equation (7.4) can be rewritten as

$$R_k = \max_{0\leq\tau<k}\Big\{(k-\tau)\Big[\hat{\lambda}_{1,\tau,k}\ln\Big(\frac{\hat{\lambda}_{1,\tau,k}}{\lambda_0}\Big)-(\hat{\lambda}_{1,\tau,k}-\lambda_0)\Big]\Big\}. \tag{7.5}$$

We see that R_k is maximized over $0 \leq \tau < k$ at each sampling point, if the sample size k is very large, the calculation of R_k becomes burdensome. In this sense, an alternative representation of R_k but without loss of much inherent information is desired. One option here is to use a window with length m, over which the maximization is taken within the past m samples, see reference Willsky *et al.* [37]. We then have

$$R_{m,k} = \begin{cases} \max\limits_{0\leq\tau<k}\Big\{(k-\tau)\Big[\hat{\lambda}_{1,\tau,k}\ln\Big(\frac{\hat{\lambda}_{1,\tau,k}}{\lambda_0}\Big)-(\hat{\lambda}_{1,\tau,k}-\lambda_0)\Big]\Big\} & \text{if } k=1,2,\cdots,m \\ \max\limits_{k-m\leq\tau<k}\Big\{(k-\tau)\Big[\hat{\lambda}_{1,\tau,k}\ln\Big(\frac{\hat{\lambda}_{1,\tau,k}}{\lambda_0}\Big)-(\hat{\lambda}_{1,\tau,k}-\lambda_0)\Big]\Big\} & \text{if } k=m+1,m+2,\cdots. \end{cases} \tag{7.6}$$

Since the R_k is based on the alternative hypothesis $H_1 : \lambda_1 \neq \lambda_0$, the value of $R_{m,k}$ is the surveillance statistic of the two-sided GLR chart. Given a control limit h_{GLR} which is designed to provide the specified in-control performance, if $R_{m,k} > h_{GLR}$, a signal will be triggered at sample k. Note that if λ_0 is small, it will be possible that $\hat{\lambda}_{1,\tau,k}$ is zero, which makes R_k in Equation (7.6) undefined. This issue will be discussed later in detail.

Let $\hat{\tau}_k$ be the value of τ at which the maximum is attained in $R_{m,k}$ in Equation (7.6), so the surveillance statistics for two-sided GLR chart at sampling point $\hat{\tau}_k$ is

$$R_{m,k} = (k-\hat{\tau}_k)\Big[\hat{\lambda}_{1,\hat{\tau}_k,k}\ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big)-(\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0)\Big]. \tag{7.7}$$

The control statistic $R_{m,k}$ in Equation (7.7) includes two important pieces of process information: the estimated changepoint $\hat{\tau}_k$ and the process mean estimator $\hat{\lambda}_{1,\hat{\tau}_k,k}$ at $\hat{\tau}_k$. In other words, if a signal is given by the GLR chart, one can get both the estimated time

point where the special cause may have occured and the magnitude of the shifted mean.

7.2 Properties of GLR chart statistics

Getting insight into features of the GLR statistic may help one generate the desired charts and take full advantage of them. Before plotting the GLR charts for Poisson data, we want to investigate the inherent nature of R_k first.

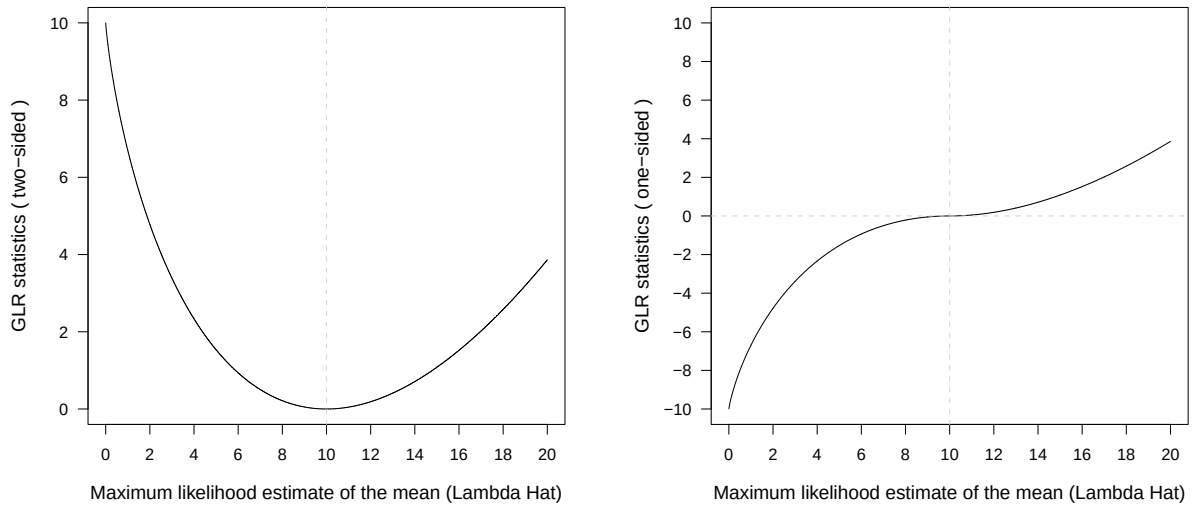
To explore the properties of R_k, consider a special case of the GLR statistic where $m = 1$. For this case, only the current sample is used, so $\hat{\tau} = k - 1$, and Equation (7.7) reduces to

$$R_{1,k} = x_k \ln(\frac{x_k}{\lambda_0}) - (x_k - \lambda_0). \tag{7.8}$$

It is not easy to fully understand the relationship between $R_{1,k}$ and x_k, since Equation (7.8) is a transcendental equation. To get insight into this equation, we plot $R_{1,k}$ as a function of x_k. Supposing $\lambda_0 = 10$, Figure 7.1(a) shows the relationship in Equation (7.8). We find that $R_{1,k}$ is an increasing function of x_k when $x_k > \lambda_0$, while $R_{1,k}$ is a decreasing function of x_k when $x_k < \lambda_0$. Also, the $R_{1,k}$ minimum value is reached at $x_k = \lambda_0$. Figure 7.1(a) also shows that $R_{1,k}$ is a two-sided control chart statistic, since any deviation from λ_0 leads to an increasing value of $R_{1,k}$.

Besides the graphical exploration, we can also investigate the properties of $R_{1,k}$ mathematically by considering the first derivative of $R_{1,k}$ with respect to x_k, i.e.,

$$\begin{aligned} \frac{\mathrm{d}R_{1,k}}{\mathrm{d}x_k} &= \ln\Big(\frac{x_k}{\lambda_0}\Big) + x_k \frac{\lambda_0}{x_k}\frac{1}{\lambda_0} - 1 \\ &= \ln\Big(\frac{x_k}{\lambda_0}\Big), \end{aligned} \tag{7.9}$$



(a) **(b)**

Figure 7.1: Relationship between GLR statistics and the maximum likelihood estimators of the mean, $\lambda_0 = 10$. (a) for the two-sided GLR; (b) for the one-sided GLR.

so we can obtain

$$\frac{\mathrm{d}R_{1,k}}{\mathrm{d}x_k} = \begin{cases} > 0 & \text{if} \quad x_k > \lambda_0 \\ = 0 & \text{if} \quad x_k = \lambda_0 \\ < 0 & \text{if} \quad x_k < \lambda_0. \end{cases} \tag{7.10}$$

Equation (7.10) gives the same information about the monotonicity of $R_{1,k}$ as shown in Figure 7.1 (a). In fact, when only the current sample is considered, the use of the GLR chart statistic is equivalent to the use of the Shewhart chart statistic. In this case, the surveillance performance of the GLR chart and the Shewhart chart are same.

Information given by Equation (7.7), the function (7.10) and Figure 7.1(a) indicates that if a signal is triggered based on the statistic $R_{m,k}$, we cannot identify whether this signal is the indicator of an increase in the mean or a decrease in the mean.

7.3 Statistics of one-sided GLR charts

In this section, we discuss how to identify the direction of the shift in the mean using the information given by Equation (7.10). To achieve this goal, one option is to use a mathematical sign function.

In mathematics, the sign function is an odd mathematical function that extracts the sign of a real number. The sign function of a real number x is defined as follows:

$$\text{sgn}(x) = \begin{cases} 1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ -1 & \text{if} \quad x < 0. \end{cases} \tag{7.11}$$

It is easy to show that any real number can be expressed as the product of its absolute value and its sign function. That is

$$x = \text{sgn}(x) \cdot |x|. \tag{7.12}$$

As discussed in section 7.1, $R_{m,k}$, $k = 1, 2, 3, \cdots$, are the statistics of a two-sided GLR chart. The application of a sign function on $R_{m,k}$ can change it from a two-sided statistic to a one-sided statistic. Let $R^s_{m,k}$ be the product of $R_{m,k}$ and the sign function of $(\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0)$, so that

$$\begin{aligned} R^s_{m,k} &= \text{sgn}(\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0) \cdot R_{m,k} \\ &= \text{sgn}(\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0) \cdot (k - \hat{\tau}_k) \Big[\hat{\lambda}_{1,\hat{\tau}_k,k} \ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0} \Big) - (\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0) \Big]. \end{aligned} \tag{7.13}$$

Figure 7.1(b) is the plot of $R^s_{m,k}$ versus $\hat{\lambda}_{1,\hat{\tau}_k,k}$ corresponding to the case illustrated

in Figure 7.1(a). The value of $R^s_{m,k}$ keeps increasing over the range $[0, 20]$. As shown in Figure 7.1(b) and Equation (7.13), the two-sided statistic $R_{m,k}$ and one-sided statistic $R^s_{m,k}$ are related; they are equivalent when $\hat{\lambda}_{1,\hat{\tau}_k,k} \geq \lambda_0$, and opposite in sign otherwise.

It can be shown that $R^s_{m,k}$ is an increasing function of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ over the range $(0, \infty)$ (see Appendix A for the proof). Because $\hat{\lambda}_{1,\hat{\tau}_k,k}$ is the estimator of the process mean at sample k, the monotonicity property of the function (7.13) allows $R^s_{m,k}$ to identify both the shift direction and the size of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ from λ_0. In particular, a positive value of $R^s_{m,k}$ means a upward shift from the target mean, whereas a negative value of $R^s_{m,k}$ means a downward shift from the target mean; the absolute value of $R^s_{m,k}$ reflects the shift size.

Note that it is possible that $\hat{\lambda}_{1,\hat{\tau}_k,k}$ is zero when λ_0 is small, which leads to Equations (7.5) and (7.13) being undefined. In this case, we can assign $\hat{\lambda}_{1,\hat{\tau}_k,k}$ a very small value (say 0.0001), which should have no effect on the chart performance, particularly if detecting an upward-shift is of interest, because when $\hat{\lambda}_{1,\hat{\tau}_k,k} \to 0$, we can conclude $R^s_{m,k} \to -\infty$ by Appendix A.

The one-sided GLR chart can work for the two-sided surveillance purpose. If the control limit of the two-sided GLR chart (h_{GLR}) has been simulated to provide a specified ATS_0, which means a signal will be triggered if $R_{m,k} > h_{GLR}$, then the corresponding one-sided GLR chart will have the same surveillance performance if a signal is given when $R^s_{m,k}$ falls outside of the interval $[-h_{GLR}, h_{GLR}]$.

Applying the sign function to $R_{m,k}$ should lead to no additional computational work load for the programming and calculation, because most software packages, such as SAS, R, and Matlab, have the built-in sign function. As a result of the benefits of using $R^s_{m,k}$ discussed above, it is recommended to use $R^s_{m,k}$ as the control statistic for both one-sided and two-sided GLR charts.

7.4 One-sided GLR charts with a reflecting barrier

There may be other approaches applied to $R_{m,k}$ to modify it to a one-sided GLR statistic, such as the use of the reflection of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ with the barrier λ_0 as described as follows:

$$\hat{\lambda}^{r+}_{1,\hat{\tau}_k,k} = \max\{\hat{\lambda}_{1,\hat{\tau}_k,k}, \lambda_0\} \tag{7.14}$$

$$\hat{\lambda}^{r-}_{1,\hat{\tau}_k,k} = \min\{\hat{\lambda}_{1,\hat{\tau}_k,k}, \lambda_0\}, \tag{7.15}$$

where $\hat{\lambda}^{r+}_{1,\hat{\tau}_k,k}$ and $\hat{\lambda}^{r-}_{1,\hat{\tau}_k,k}$ are upward-sided and downward-sided parameter estimators reflected with the barrier λ_0, respectively. Thus, the upward-sided statistic ($R^{r+}_{m,k}$) and downward-sided statistic ($R^{r-}_{m,k}$) of the GLR chart can be expressed in the forms of Equation (7.16) and Equation (7.17), respectively, given by

$$R^{r+}_{m,k} = (k - \hat{\tau}_k)\Big[\hat{\lambda}^{r+}_{1,\hat{\tau}_k,k}\ln\Big(\frac{\hat{\lambda}^{r+}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big) - (\hat{\lambda}^{r+}_{1,\hat{\tau}_k,k} - \lambda_0)\Big] \tag{7.16}$$

$$R^{r-}_{m,k} = (k - \hat{\tau}_k)\Big[\hat{\lambda}^{r-}_{1,\hat{\tau}_k,k}\ln\Big(\frac{\hat{\lambda}^{r-}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big) - (\hat{\lambda}^{r-}_{1,\hat{\tau}_k,k} - \lambda_0)\Big]. \tag{7.17}$$

Taking the upward-sided statistic $R^{r+}_{m,k}$ as an example, if $\hat{\lambda}_{1,\hat{\tau}_k,k}$ is less than λ_0, then $\hat{\lambda}_{1,\hat{\tau}_k,k}$ is set to λ_0 by the reflection approach, so $R_{m,k} = 0$ if $\hat{\lambda}_{1,\hat{\tau}_k,k} < \lambda_0$ as shown in Figure 7.2.

Although $R^{s}_{m,k}$, $R^{r+}_{m,k}$ and $R^{r-}_{m,k}$ are represented in different forms, all of them are step functions of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ mathematically. To develop the traditional upward-sided, downward-sided and two-sided GLR charts, the reflection approach requires one to calculate three surveillance statistics using Equations (7.16), (7.17), and (7.7), respectively; whereas the sign function approach has the uniform expression given in Equation (7.13) for three surveillance statistics. As a result, the users of GLR charts can benefit from the unique and concise feature of the sign function applied approach in practice. The performance of the GLR chart is the same regardless of which of the two approaches is taken.

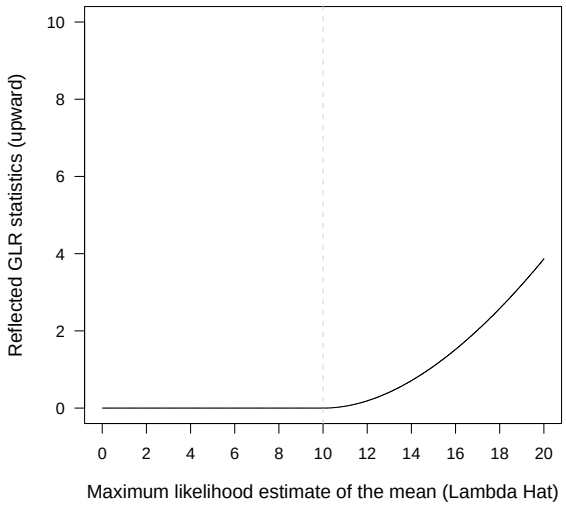


Figure 7.2: Relationship between reflected upward-sided GLR statistics and the maximum estimator of the mean when $\lambda_0 = 10$.

Chapter 8

Design of GLR control charts

In the previous chapter, we derived the GLR statistics for both one-sided and two-sided charts. The use of GLR charts involves some pre-determined settings such as the window length and the control limit. Also, due to the complexity of the statistics $R^s_{m,k}$, more detailed guidance will be very useful for the GLR chart users. Considering the issues involving in the use of the GLR chart, in this chapter we first discuss how to determine the appropriate window length and the control limit, and then provide guidance for the GLR statistics calculation and the GLR chart plotting with some illustrative examples.

8.1 Choosing the window length

The use of a window to calculate the GLR statistic can ease the burdensome calculation as discussed in section 7.1. In this section, we will discuss how to choose a reasonable window length. The simulation of ATS_0 for the GLR chart is time-consuming work, as mentioned in other GLR literature [44, 47]. Moreover, the use of either proportional shifts or standardized shifts, fortunately, should reach similar conclusions on the relative performance of the competing methods due to the relationship of these two shifts.

In order to explore the effect of varying window lengths on the CED performance, Table 8.1 summarizes CED values for a combination of ranges of m and δ for two-sided GLR charts with $\lambda_0 = 2$. The first row, labeled by m, lists the choices of m ranging from the smallest possible value of 1 to the relatively large value of 1200; the last row in Table 8.1, labeled by h_{GLR}, lists the control limits corresponding to the desired $ATS_0 = 1500$ for each m ; the first column lists the choices of δ ranging from 0 to 7, where $\delta = 0$ is the in-control condition and the CED is consequently ATS_0.

Table 8.1: Effect of the choice of m on h_{GLR} and on the CED of the two-sided GLR chart for sustained shifts in the process with $\lambda_0 = 2$

$m =$	1	5	10	25	50	100	200	400	800	1200
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
0.00	4149.86	1826.07	1530.62	1503.24	1553.37	1515.80	1500.05	1500.48	1502.02	1500.14
0.25	885.84	379.51	246.58	186.62	137.20	104.33	93.29	91.75	91.61	91.57
0.50	262.84	77.06	45.53	32.83	28.59	27.74	27.80	27.78	27.79	28.04
0.75	100.71	25.32	16.30	14.30	14.16	14.13	14.16	14.15	14.18	14.25
1.00	46.42	11.87	8.84	8.78	8.88	8.87	8.87	8.86	8.88	8.87
1.50	14.34	4.69	4.31	4.56	4.64	4.63	4.63	4.63	4.63	4.63
2.00	6.13	2.78	2.78	2.90	2.99	2.99	2.98	2.98	2.98	2.98
3.00	1.98	1.51	1.53	1.54	1.61	1.61	1.61	1.61	1.61	1.62
4.00	0.99	0.99	1.00	1.00	1.03	1.03	1.03	1.03	1.03	1.04
5.00	0.68	0.73	0.73	0.73	0.74	0.74	0.74	0.74	0.74	0.75
6.00	0.56	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61
7.00	0.52	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54
h_{GLR}	5.0912	6.6975	6.7762	7.3394	7.5396	7.5402	7.5629	7.5780	7.5816	7.6163

Although the desired ATS_0 value for each m is 1500, it is impossible to generate the exact value due to the simulation errors and the discrete nature of the Poisson distribution. In our simulations, we set ATS_0 as close to 1500 as possible without going below it. For $m = 1$ (labeled [1]), for example, the simulated ATS_0 is 4149.86, which is far away from 1500. This extreme departure from the desired value should not be surprising, because of the discreteness for the Poisson data. Furthermore, the smaller mean in this case, $\lambda_0 = 2$, strengthens that departure. We also note that, compared to using a larger m, the use of a smaller m value makes it harder to approach 1500, which results from the fact that GLR charts with a smaller m use less past information than those with a larger m value.

The GLR chart with $m = 1$ is equivalent to the Shewhart chart, since both of them only use the current sample. See Appendix B for the theoretical proof for their equivalence. This fact will be discussed in more detail in later sections. Table 8.1 indicates that the GLR chart with $m = 1$ has poorer performance for detecting small mean shifts than for detecting large mean shifts, which is similar to the usual Shewhart chart performance.

We can see that h_{GLR} increases as m increases. As indicated in Equation (7.6), the GLR statistic $R_{m,k}$ is maximized over all samples in a window with length m. The larger number of m means $R_{m,k}$ is maximized over a larger number of past samples, so the GLR statistic with larger m will consequently never be less than that with smaller value of m, which leads to an increase in the control limit to keep ATS_0 at the same level.

Since detecting a small shift requires a relatively large amount of information, an adequately large value of m is required to effectively catch these small shifts. From Table 8.1, we find that, in terms of CED, there are no significant differences beyond $m \geq 400$ even for the small shift $\delta = 0.25$, so it seems that $m = 400$ is large enough to detect small shifts efficiently. The performance with $m = 1200$ is likely very close to the performance without applying a window (i.e., $m = \infty$ in Equation (7.6)). When comparing the performance of $m = 400$ with $m = 5$, we also find that, even for large shifts, there is no significant difference in their CED values. Combining the information given by Table 8.1, we conclude that $m = 400$ is a suitable window length for practical use of GLR charts to detect a range of mean shifts.

8.2 Choosing the control limits

If a reasonable window length has been determined, then the only parameter for the GLR chart is the control limit, which is determined to give a specified in-control performance, i.e., ATS_0. The investigation of the window length in the preceding section indicates that $m = 400$ is a quite reasonable window length in practice. In this section, the determination of control limits is discussed.

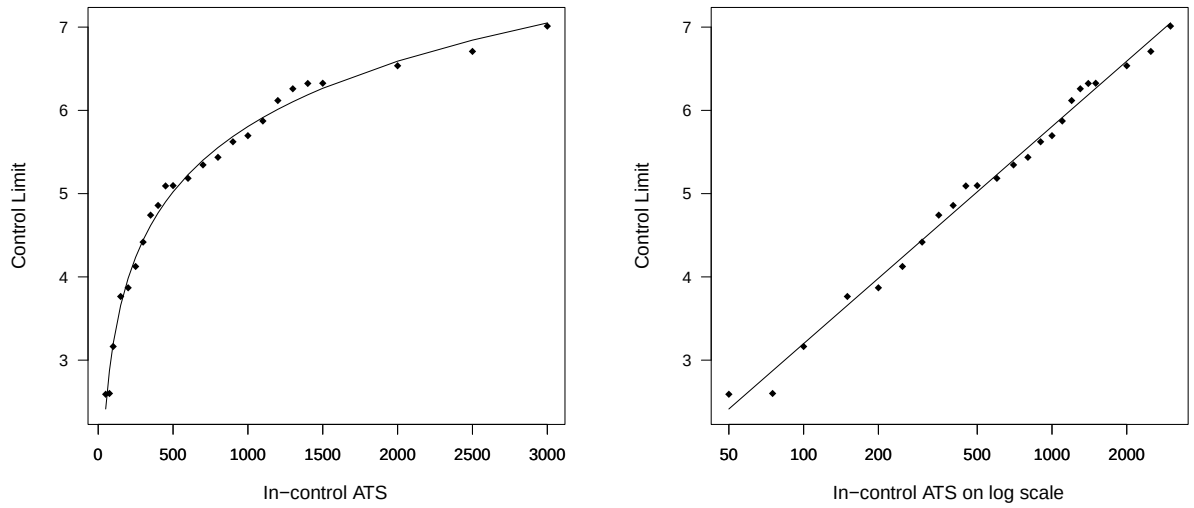
For the one-sided GLR chart with $\lambda_0 = 2$ and $m = 400$, Table 8.2 gives the control limits for some ATS_0 values ranging from 50 to 3000, as well as CED values for the combination of ATS_0 and δ within a range $[0.25, 7]$. As discussed in Section 6.1, some desired ATS_0 values for Poisson data cannot be exactly achieved due to the discreteness of Poisson data. For the purpose of simplicity, the desired ATS_0 values are displayed in Table 8.2 , instead of the ATS_0 values determined by simulation. Simulated ATS_0 values based on the control limits listed in Table 8.2 should be the closest ATS_0 values to the desired ones.

As expected, Table 8.2 shows that the CED decreases as δ increases, and that h_{GLR} increases as ATS_0 increases. To illustrate how h_{GLR} changes with ATS_0 intuitively, Figure 8.1(a) gives the plot of h_{GLR} as a function of ATS_0. It seems that an exponential function can describe their relationship approximately. To confirm this, h_{GLR} against ATS_0 on the log scale is plotted in Figure 8.1(b), in which a straight line indicates that there is a linear relationship between h_{GLR} and $\log(ATS_0)$.

This linear relationship makes it easier to achieve a specific ATS_0. For example, to generate Table 8.2, we can start with two plausible values of h_{GLR}, then using simulations we generate the two corresponding values of ATS_0, so we can get the linear function of h_{GLR} and $\log(ATS_0)$. Thus, other values of h_{GLR} can be estimated by this linear relationship.

Table 8.2: Values of h_{GLR} corresponding to specified values of the in-control ATS_0 and the CED for sustained shifts for the one-sided GLR chart with $m = 400$ and $\lambda_0 = 2$

		CED										
ATS_0	h_{GLR}	δ=0.25	0.50	0.75	1.00	1.50	2.00	3.00	4.00	5.00	6.00	7.00
50	2.5903	10.93	7.09	4.70	3.44	2.10	1.49	0.92	0.69	0.59	0.54	0.52
75	2.5998	13.05	8.16	5.28	3.82	2.35	1.66	1.05	0.77	0.64	0.57	0.53
100	3.1639	18.68	11.21	7.13	4.99	2.96	2.04	1.21	0.85	0.67	0.58	0.54
150	3.7646	24.56	14.02	8.77	5.91	3.37	2.25	1.28	0.88	0.69	0.59	0.54
200	3.8690	29.18	16.12	9.89	6.66	3.83	2.56	1.48	1.02	0.78	0.65	0.58
250	4.1251	43.55	19.14	11.01	7.24	4.02	2.66	1.51	1.03	0.78	0.65	0.58
300	4.4178	51.11	21.57	12.19	7.95	4.35	2.87	1.62	1.09	0.81	0.66	0.59
400	4.8582	62.36	24.79	13.88	8.99	4.87	3.18	1.73	1.13	0.82	0.67	0.59
500	5.0964	71.61	27.14	14.88	9.59	5.15	3.37	1.86	1.27	0.94	0.75	0.64
600	5.1838	76.17	28.42	15.47	9.98	5.37	3.52	1.98	1.34	0.98	0.76	0.65
700	5.3449	79.39	29.37	15.87	10.17	5.42	3.54	1.98	1.34	0.98	0.76	0.65
800	5.4362	84.34	30.77	16.56	10.65	5.73	3.72	2.04	1.36	0.98	0.77	0.65
900	5.6228	88.63	31.97	17.11	10.95	5.84	3.77	2.05	1.36	0.99	0.77	0.65
1000	5.6962	92.70	33.05	17.67	11.28	6.03	3.90	2.12	1.39	1.00	0.77	0.65
1100	5.8722	96.51	34.20	18.16	11.54	6.11	3.94	2.13	1.39	1.00	0.77	0.65
1200	6.1180	108.85	36.22	18.94	11.94	6.24	3.99	2.14	1.40	1.00	0.78	0.65
1300	6.2593	113.92	37.51	19.58	12.38	6.46	4.11	2.16	1.40	1.01	0.78	0.65
1400	6.3232	117.95	38.39	20.02	12.65	6.62	4.23	2.27	1.47	1.04	0.79	0.66
1500	6.3259	119.55	38.74	20.18	12.77	6.73	4.32	2.30	1.48	1.05	0.79	0.66
2000	6.5357	125.73	40.36	20.91	13.12	6.85	4.36	2.31	1.48	1.05	0.79	0.66
2500	6.7081	137.42	42.54	21.78	13.68	7.15	4.60	2.52	1.65	1.19	0.91	0.74
3000	7.0120	148.97	45.56	23.07	14.47	7.58	4.82	2.57	1.66	1.20	0.92	0.74



(a) **(b)**

Figure 8.1: The GLR control limit h_{GLR} of the GLR chart with $m = 400$ as a function of the in-control limits. (a) with the original ATS scale; (b) with log(ATS) scale.

Table 8.2 lists some practical values of h_{GLR} for health-care surveillance, but other other values of h_{GLR} can be approximated based on this table. For example, if the value of h_{GLR} for $\text{ATS}_0 = 350$ is desired, we can estimate the h_{GLR} by using the paired values of $(300, 4.4178)$ and $(400, 4.8582)$ as follows:

$$\begin{aligned} h_{GLR} &= 4.4178 + (4.8582 - 4.4178)\frac{\ln(400) - \ln(300)}{\ln(350) - \ln(300)} \\ &= 4.6538. \end{aligned} \tag{8.1}$$

Based on simulation, the actual value of ATS_0 is 339.47 for $h_{GLR} = 4.6538$. Considering the discreteness of Poisson data as discussed earlier, it is reasonable to say that 339.47 is close to the desired value of 350.

8.3 Example calculation for GLR charts

The one-sided GLR chart is a plot of $R^s_{m,k}$ versus k. Given a sample, the value of $R^s_{m,k}$ can be calculated by using Equations (7.3),(7.6),(7.11) and (7.13), but it may be not so clear how to compute these values. In order to clarify the procedures involved, we summarize the steps for calculating $R^s_{m,k}$ as the following:

Given a sample $\{X_1, X_2, \cdots, X_k\}$, for the purpose of demonstration, we take the case with the use of a window as an example here, so we assume $k > m$.

1. Start with $\tau = k - m$, calculate $\hat{\lambda}_{1,\tau,k} = \frac{\sum_{i=\tau+1}^{k}}{k-\tau}$;
2. Calculate $\hat{\beta}_{\tau,k} = (k-\tau)\Big[\hat{\lambda}_{1,\tau,k}\ln\Big(\frac{\hat{\lambda}_{1,\tau,k}}{\lambda_0}\Big) - (\hat{\lambda}_{1,\tau,k} - \lambda_0)\Big]$;
3. Calculate a value using the sign function

$$\text{sgn}(\hat{\lambda}_{1,\tau,k} - \lambda_0) = \begin{cases} 1 & \text{if } \hat{\lambda}_{1,\tau,k} > \lambda_0 \\ 0 & \text{if } \hat{\lambda}_{1,\tau,k} = \lambda_0 \\ -1 & \text{if } \hat{\lambda}_{1,\tau,k} < \lambda_0; \end{cases}$$

4. Calculate $\hat{\beta}^s_{\tau,k} = \hat{\beta}_{\tau,k} \cdot \text{sgn}(\hat{\lambda}_{1,\tau,k} - \lambda_0)$;
5. Repeat above steps (1),(2),(3) and (4) for $\tau = (k-m+1), (k-m+2), \cdots, (k-1)$; so we get $\hat{\beta}^s_{k-m,k}, \hat{\beta}^s_{k-m+1,k}, \cdots, \hat{\beta}^s_{k-1,k}$;
6. Calculate $R^s_{m,k} = \max\{\hat{\beta}^s_{k-m,k}, \hat{\beta}^s_{k-m+1,k}, \cdots, \hat{\beta}^s_{k-1,k}\}$;
7. Suppose $R^s_{m,k} = \hat{\beta}^s_{\hat{\tau}_k,k}$ in step (6). This means the estimated changepoint by the GLR algorithm for the sample k is $\hat{\tau}_k$, so the estimated process mean by the GLR algorithm for the sample k is $\hat{\lambda}_{1,\hat{\tau}_k,k}$.

An example of the calculation is illustrated in Table 8.3. The first column, labeled by k, is the observation index; the second column, labeled by x_k, is the observation value; the column $\hat{\tau}_k$ is the index of τ at which $R^s_{m,k}$ is obtained; two-sided GLR statistics are listed in the column $R_{m,k}$, which contains the corresponding absolute values of $R^s_{m,k}$.

Table 8.3: Example of the calculation of the statistics needed in applying the GLR chart when $m = 3$ and $\lambda_0 = 2$

k	x_k	$\hat{\lambda}_{1,k-3,k}$	$\hat{\lambda}_{1,k-2,k}$	$\hat{\lambda}_{1,k-1,k}$	$\hat{\beta}_{k-3,k}$	$\hat{\beta}_{k-2,k}$	$\hat{\beta}_{k-1,k}$	$\hat{\tau}_k$	$\hat{\lambda}_{1,\hat{\tau}_k,k}$	$R_{m,k}$	$R^s_{m,k}$
1	1	—	—	1.000	—	—	0.307	0	1.000	0.307	-0.307
2	2	—	1.500	2.000	—	0.137	0.000	0	1.500	0.137	-0.137
3	5	2.667	3.500	5.000	0.301	0.917	1.581	2	5.000	1.581	1.581
4	2	3.000	3.500	2.000	0.649	0.917	0.000	2	3.500	0.917	0.917
5	5	4.000	3.500	5.000	2.318	0.917	1.581	2	4.000	2.318	2.318
6	2	3.000	3.500	2.000	0.649	0.917	0.000	4	3.500	0.917	0.917
7	3	3.333	2.500	3.000	1.108	0.116	0.216	4	3.333	1.108	1.108
8	6	3.667	4.500	6.000	1.667	2.298	2.592	7	6.000	2.592	2.592
9	9	6.000	7.500	9.000	7.775	8.826	6.537	7	7.500	8.826	8.826
10	5	6.667	7.000	5.000	10.080	7.539	1.581	7	6.667	10.080	10.080

For the purpose of illustration, the data in Table 8.3 with $\lambda_0 = 2$ is artificially constructed, and a small window length $m = 3$ is assumed in this example. It can be seen that there are some big jumps between adjacent data values, which is designed to let one easily check how the value of $R^s_{m,k}$ responds to these big changes. Because no window will be used if $k < m$, some values in Table 8.3 should be missing. When the observed values suddenly increase from 2 (observation 2) to 5 (observation 3), the GLR statistic $R^s_{m,k}$ changes from -0.0137 to 1.581, which indicates that $R^s_{m,k}$ has a quick response to the change in observations.

As discussed earlier, the GLR algorithm has the inherent ability to estimate the changepoint ($\hat{\tau}_k$) and the estimated process mean ($\hat{\lambda}_{1,\hat{\tau}_k,k}$). We can see that the last three values of $\hat{\tau}_k$ are 7, which indicates that the process mean may have been changed from the target mean at sampling point 7, and we also see that the estimated mean $\hat{\lambda}_{1,\hat{\tau}_k,k}$ is

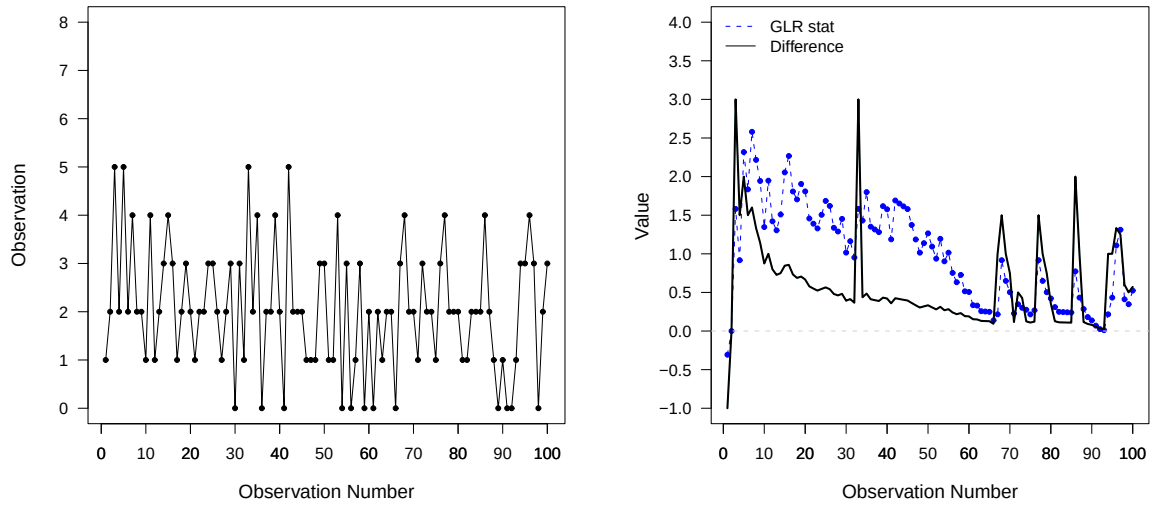
larger than the previous values and the last three values of $R^s_{m,k}$ increase suddenly. This table implies $R^s_{m,k}$ has quickly responded to the increased observations. Whether a signal will be given here depends on the h_{GLR} value, which is determined to provide desired in-control chart performance.

8.4 Example plots of GLR charts

One-sided GLR control charts are plots of $R^s_{m,k}$ against k, while two-sided GLR control charts are plots of $R_{m,k}$ against k. In this section, we will illustrate what one-sided GLR charts look like and discuss some of the chart properties.

Figure 8.2(a) is a plot of one hundred in-control observations for Poisson data with $\lambda_0 = 2$. Since the mean is 2, it is reasonable that some observations are zeros. In Figure 8.2(b), the dashed line represents the values of $R^s_{m,k}$ shown in Figure 8.2(a), and the solid line represents the values of the difference between $\hat{\lambda}_{1,\hat{\tau}_k,k}$ and λ_0, i.e., $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$, which is a measure of how far the process mean departs from λ_0 using the maximum likelihood estimator $\hat{\lambda}_{1,\hat{\tau}_k,k}$. Note that the scales of $R^s_{m,k}$ and $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$ are not the same, but plotting them in one graph should have no effect on the comparison and analysis.

Figure 8.3 is a plot of the two-sided statistics of GLR charts (i.e., $R_{m,k}$) versus their observation numbers. We can see that neither Figure 8.3 nor Figure 8.2(b) looks like traditional two-sided control charts for the mean, for which the surveillance statistics are close to a center line when a process is in control. As discussed in Section 7.2, the statistics $R_{m,k}$ should never be negative, and Figure 8.3 displays that they are typically not close to a center line even when the process in control. In contrast, the two-sided GLR chart looks more like the control charts for detecting an increase in the standard deviation.



(a) **(b)**

Figure 8.2: Plot of in-control observations, $R^s_{m,k}$ and $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$ for Poisson data with $\lambda_0 = 2$. (a) the plot of in-control observations; (b) the dashed line (GLR stat) represents $R^s_{m,k}$, and the solid line (Difference) represents $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$.

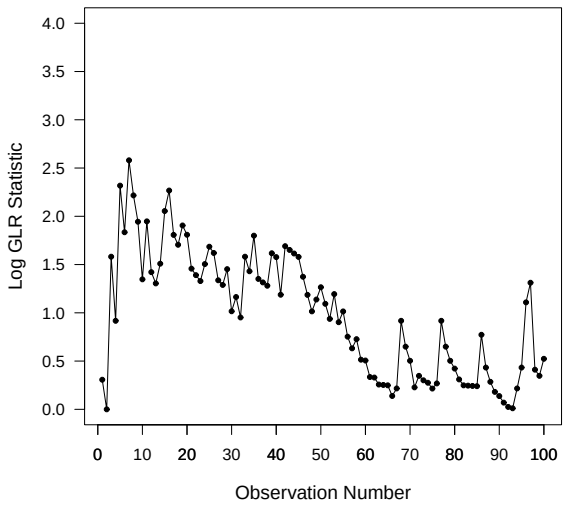
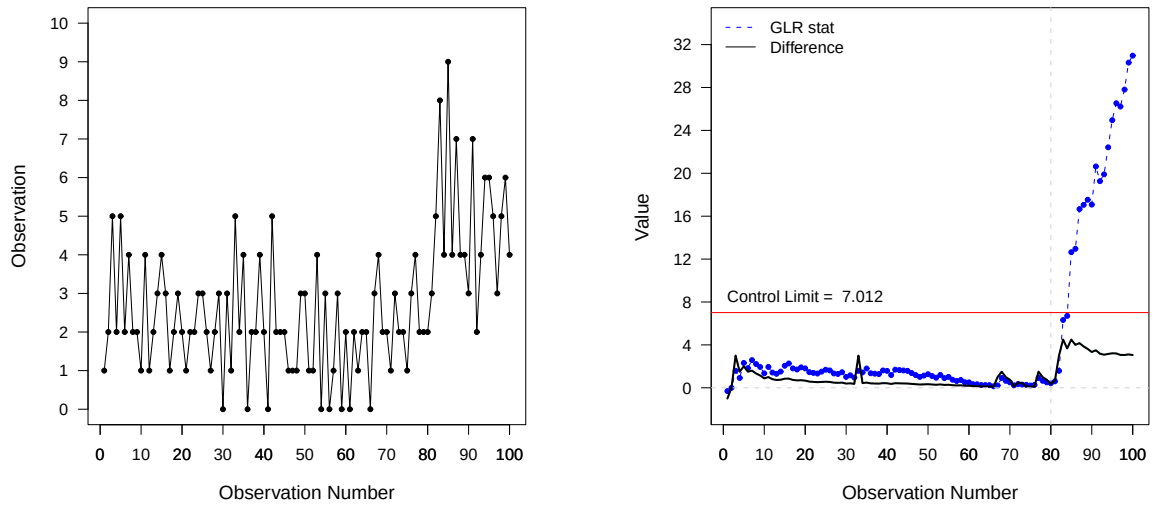


Figure 8.3: Plot of $R_{m,k}$ for the in-control observations when $\lambda_0 = 2$.

If a sample $\{X_1, X_2, \cdots, X_k\}$ is taken, based on the Equation (7.3), the current value of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ will depend on τ that maximizes the likelihood-ratio statistic. In fact, there may be significant changes in the observed values from one observation to the next, so the value of $\hat{\tau}_k$ may vary dramatically, which leads to significant changes in $\hat{\lambda}_{1,\hat{\tau}_k,k}$ and $R^s_{m,k}$ from one observation to the next as well. For example, in Figure 8.2(b), at time $k = 32$, $\hat{\tau}_{32} = 2$ and $\hat{\lambda}_{1,2,32} - \lambda_0 = 0.37$, so $\hat{\lambda}_{1,2,32}$ is based on the past 31 observations. However, at time $k = 33$, observation $x_{33} - \lambda_0 = 3$ leads to the estimated changepoint dramatically changing to $\hat{\tau}_{33} = 32$. In the latter case, the estimate of λ_1 is only based on the last observation and this results in $\hat{\lambda}_{1,32,33} - \lambda_0 = 3$. When the process is in control, the large variation in $\hat{\tau}_k$ and $\hat{\lambda}_{1,\hat{\tau}_k,k}$ result from the in-control observations, which are random during the simulation.

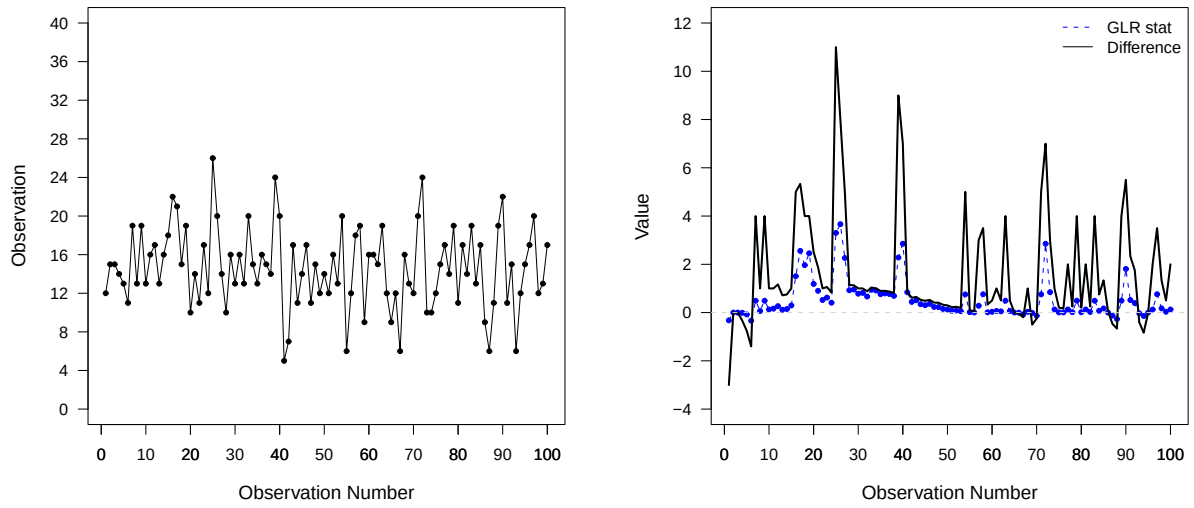
Figure 8.4(a) is a plot of observations in which a mean shift with $\delta = 1.5$ occurs after $\tau = 80$. Otherwise it is the same as Figure 8.2(a). Figure 8.4(b) is the corresponding plot of $R^s_{m,k}$ and $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$. With the control limit $h_{GLR} = 7.012$, this shift is detected by the GLR chart at observation x_{85}. At the time of the signal, $\hat{\tau}_{85} = 81$ and $\hat{\lambda}_{1,81,85} - \lambda_0 = 4.5$, so the GLR chart has quickly responded to the shift, and the chart produces good estimates of the time as well as the size of the shift in the mean.

In addition, we also studied the case in which $\lambda_0 = 15$. For Poisson data with $\lambda_0 = 15$, Figures 8.5 and 8.6 are the plots for in-control and out-of-control cases, respectively. The mean shift point and m of these two charts are the same as Figures 8.2 and 8.4 except $\lambda_0 = 15$ and $h_{GLR} = 7.254$. The features represented in Figures 8.5 and 8.6 are similar to those shown in Figures 8.2 and 8.4.



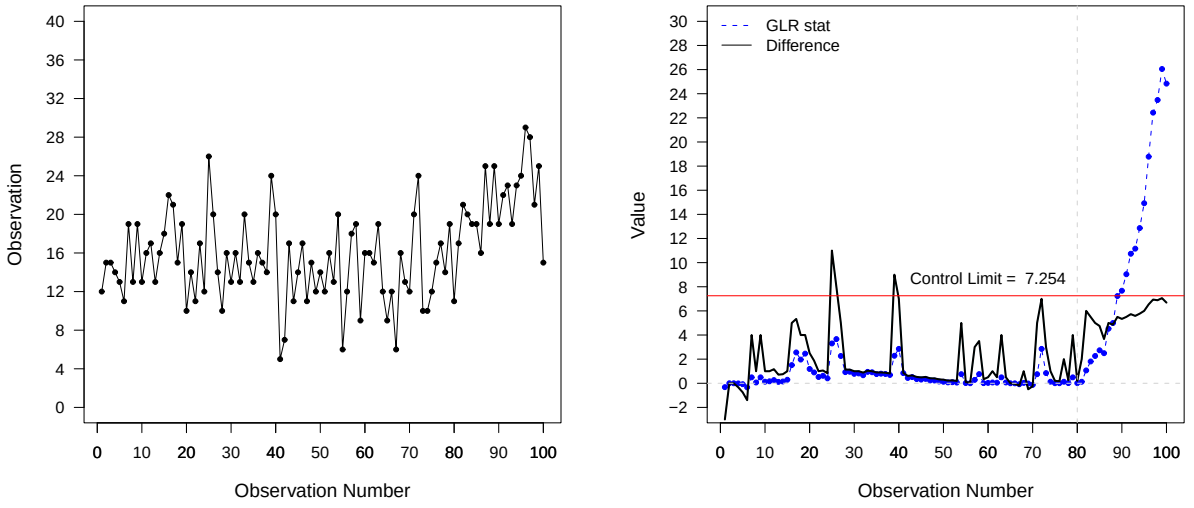
(a) **(b)**

Figure 8.4: Plot of out-control observations, $R^s_{m,k}$ and $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$ for Poisson data with $\lambda_0 = 2$, $\delta = 1.5$ and $\tau = 80$. (a) plot of out-control observations; (b) the dashed line (GLR stat) represents $R^s_{m,k}$, and the solid line (Difference) represents $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$.



(a) (b)

Figure 8.5: Plot of in-control observations, $R^s_{m,k}$ and $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$ for Poisson data with $\lambda_0 = 15$. (a) the plot of in-control observations; (b) the dashed line (GLR stat) represents $R^s_{m,k}$, and the solid line (Difference) represents $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$.



(c) (d)

Figure 8.6: Plot of out-control observations, $R^s_{m,k}$ and $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$ for Poisson data with $\lambda_0 = 15$, $\delta = 1.5$ and $\tau = 80$. (c) the plot of out-control observations; (d) the dashed line (GLR stat) is for $R^s_{m,k}$, and the solid line (Difference) is for $\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0$.

Chapter 9

Performance comparisons

In this chapter, the GLR chart performance will be compared with alternative methods using a combination of various shift sizes and in-control ATS values.

Evaluation of control charts for normal data can be based without loss of generality on one simple case, the standard normal distribution [36, 38, 85, 94]. However, the chart performance for Poisson counts may vary with the magnitude of the in-control mean because of the skewness of Poisson distribution [62]. Poisson data with a small rate of $\lambda_0 = 2$ and a medium rate of $\lambda_0 = 10$ were therefore used in our simulation study.

We conducted 10,000 simulations to estimate the conditional expected delay (CED) for each chart, with each CED simulated under the steady-state conditions as discussed in Section 9. We used sustained standardized shifts in the mean with a wide range of sizes from 0.25 to 7. To explore the best setting for alterative charts, such as CUSUM and EWMA charts, various parameters associated with them were also applied.

Findings and results derived from simulations often depend on the simulation settings involved. It is necessary here to introduce the global settings used in this simulation study.

1. Simulation platform. All simulation work in this study was conducted using the software R 2.11. Compared with other simulation environments, R has an appealing feature, by which one can treat either rows or columns of a matrix as an calculation unit to compute the desired statistics, which speeds the simulations and consequently saves much time spent on the large amount of simulation work involved.

2. Number of repetitions. Most of the results in this study are based on 10,000 simulations.

3. Simulation seeds. The seeds used in this study are set as uniform as possible for different methods, which leads to make the comparisons between different approaches on the nearly same data basis.

4. Change points of time. If a simulation involves the shifts in the mean, then the change point of time will be set at time=1000. Since our study is based on the steady-state condition, if a control chart signals before the changepoint, this run was ignored and another run added. For this case, there were some differences in the seeds of different detection methods. In addition, the changepoint effect on chart performance has been tested by changing 1000 to other values; we found that there was no significant differences in terms of CED as long as the change point values were reasonably large.

The GLR chart performance was compared with alternatives in terms of CED, EQL, ERCED, and AMOC . Simulation results are generally summarized in tables, and some important results or findings are displayed in some plots.

9.1 Comparison using CED, EQL and ERCED

Table 9.1 gives CED, EQL and ERCED values of all charts studied for Poisson data with $\lambda_0 = 2$. For the GLR chart, we used $m = 400$. The tuning parameter of CUSUM, δ_1, defined by the standardized shift $(\lambda - \lambda_0)/\sqrt{\lambda_0}$, ranges from the small value of 0.25 to a relatively large value of 3; the smoothing parameter of EWMA, α, ranges from 0.05 to 0.4 with 5 options. The sustained mean shift within a range of [0.25, 7] is listed at the beginning of some rows. The desired ATS_0 is 1500, but there is some deviation from this value for some cases(e.g., column [11]) due to the simulation error or the discrete nature of the Poisson data. The values in the row starting with $\delta = 0.00$ lists the simulated ATS_0 for the corresponding charts, and control limits are displayed at the bottom of this table.

As mentioned earlier, we designed the charts to set ATS_0 as close as 1500 as possible without going below it in our simulation study. However, an exception can be found for the Shewhart chart, which has two values of ATS_0, 921.45 in column [3] and 4150.73 in column [4]. Checking with the simulations, we find there are no additional values of ATS_0 between these two values. This should not surprise us because of the discreteness of Poisson data as discussed in Section 6.1.

We see that for $\delta < 3$, the CED values of the EARS method in column [2] are more than 500; even for a large value of $\delta = 7$, its CED is still as large as 42. The CED values of EARS are extremely large compared with those of other charts across the entire shift range, so we will not consider it further in this study. In fact, if the EARS method does not detect a shift within 7 time points, it does not work well, since the baseline will begin to include the outbreak data.

The Shewhart chart with $ATS_0 = 921.45$ performs well for large shifts, but performs not well for small shifts in terms of CED values. For example, the Shewhart chart is

Table 9.1: CED, EQL and ERCED values for charts with $\lambda_0 = 2$ for sustained shifts in the process mean

	GLR	EARS	Shewhart		CUSUM							EWMA				
δ_1 =	—	—	—	—	0.25	0.5	1	1.5	2	2.5	3	—	—	—	—	—
α =	—	—	—	—	—	—	—	—	—	—	—	0.05	0.1	0.2	0.3	0.4
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
0.00	1549.56	1508.38	921.45	4150.73	1501.55	1505.95	1558.83	1564.19	1684.33	1512.79	1712.84	1501.80	1501.29	1502.99	1500.27	1538.85
CED values																
0.25	119.55	1111.92	339.92	1333.99	80.81	101.92	159.15	215.82	290.69	306.47	391.41	118.07	157.08	233.35	290.83	344.10
0.50	38.74	933.11	148.42	520.47	31.93	31.66	40.95	55.50	77.63	89.17	118.31	35.60	42.89	63.75	85.13	106.58
0.75	20.18	795.49	75.74	231.56	19.72	17.80	18.94	22.54	29.97	35.48	45.64	19.31	20.15	26.26	34.37	44.34
1.00	12.77	714.37	42.35	117.00	14.17	12.23	11.78	12.77	15.42	17.67	21.94	13.10	12.60	14.47	17.93	22.33
1.50	6.73	600.09	16.76	39.70	9.08	7.44	6.50	6.39	6.83	7.27	8.34	7.96	7.06	7.00	7.65	8.75
2.00	4.32	517.36	8.25	17.12	6.65	5.33	4.44	4.17	4.18	4.20	4.55	5.74	4.90	4.50	4.55	4.86
3.00	2.30	371.26	2.99	5.13	4.32	3.38	2.69	2.40	2.26	2.19	2.21	3.69	3.02	2.58	2.42	2.38
4.00	1.48	246.19	1.51	2.30	3.20	2.47	1.93	1.67	1.52	1.49	1.42	2.73	2.19	1.80	1.63	1.54
5.00	1.05	151.13	0.95	1.30	2.54	1.95	1.50	1.26	1.12	1.11	1.02	2.16	1.72	1.38	1.21	1.12
6.00	0.79	82.28	0.71	0.87	2.11	1.60	1.22	0.99	0.88	0.88	0.78	1.79	1.40	1.10	0.95	0.87
7.00	0.66	42.20	0.60	0.69	1.80	1.38	1.01	0.82	0.73	0.72	0.66	1.52	1.18	0.91	0.78	0.72
EQL(δ) values																
0.25	28.61	266.06	81.34	319.19	19.34	24.39	38.08	51.64	69.56	73.33	93.66	28.25	37.59	55.84	69.59	82.34
0.50	23.38	563.18	89.58	314.13	19.27	19.11	24.72	33.50	46.85	53.82	71.41	21.49	25.89	38.48	51.38	64.33
0.75	22.05	869.34	82.77	253.06	21.55	19.45	20.70	24.63	32.75	38.77	49.88	21.10	22.02	28.70	37.56	48.46
1.00	21.80	1219.51	72.30	199.73	24.19	20.88	20.11	21.80	26.32	30.17	37.45	22.36	21.51	24.70	30.61	38.12
1.50	22.28	1986.69	55.49	131.43	30.06	24.63	21.52	21.16	22.61	24.07	27.61	26.35	23.37	23.18	25.33	28.97
2.00	23.39	2801.10	44.67	92.69	36.01	28.86	24.04	22.58	22.63	22.74	24.64	31.08	26.53	24.36	24.64	26.31
3.00	25.58	4128.90	33.25	57.05	48.04	37.59	29.92	26.69	25.13	24.36	24.58	41.04	33.59	28.69	26.91	26.47
4.00	27.87	4635.37	28.43	43.31	60.25	46.51	36.34	31.44	28.62	28.05	26.74	51.40	41.23	33.89	30.69	29.00
5.00	29.96	4312.58	27.11	37.10	72.48	55.64	42.80	35.96	31.96	31.67	29.11	61.64	49.08	39.38	34.53	31.96
6.00	31.79	3311.16	28.57	35.01	84.91	64.39	49.10	39.84	35.41	35.41	31.39	72.03	56.34	44.27	38.23	35.01
7.00	35.61	2276.68	32.37	37.23	97.11	74.45	54.49	44.24	39.38	38.84	35.61	82.00	63.66	49.09	42.08	38.84
EQL ratios																
G(1,1)	1	65.75	2.90	8.75	1.17	1.03	1.06	1.20	1.49	1.65	2.04	1.13	1.13	1.34	1.63	1.95
G(1,2)	1	94.56	2.31	6.26	1.49	1.23	1.12	1.14	1.28	1.38	1.61	1.35	1.22	1.26	1.40	1.59
G(1,3)	1	105.25	2.03	5.19	1.66	1.34	1.16	1.13	1.21	1.28	1.45	1.47	1.29	1.25	1.33	1.46
G(2,1)	1	101.89	2.25	5.78	1.51	1.23	1.09	1.08	1.19	1.27	1.46	1.34	1.19	1.19	1.31	1.47
G(2,2)	1	125.09	1.59	3.40	1.92	1.51	1.22	1.11	1.09	1.11	1.16	1.65	1.37	1.21	1.19	1.22
G(3,1)	1	127.81	1.67	3.58	1.82	1.44	1.17	1.07	1.06	1.08	1.14	1.57	1.31	1.17	1.16	1.20
G(3,2)	1	133.41	1.21	2.09	2.19	1.70	1.32	1.14	1.06	1.05	1.02	1.87	1.51	1.25	1.14	1.11
U(.25,7)	1	119.83	1.48	3.14	2.05	1.60	1.29	1.15	1.12	1.15	1.18	1.76	1.45	1.26	1.22	1.24
ERCED ratios																
G(1,1)	1	65.31	2.93	8.79	1.17	1.02	1.05	1.19	1.47	1.63	2.03	1.12	1.12	1.32	1.62	1.94
G(1,2)	1	92.03	2.39	6.52	1.45	1.20	1.10	1.13	1.28	1.38	1.63	1.31	1.20	1.25	1.41	1.61
G(1,3)	1	102.48	2.13	5.52	1.60	1.30	1.14	1.12	1.22	1.30	1.48	1.42	1.25	1.24	1.34	1.48
G(2,1)	1	99.20	2.31	6.00	1.47	1.21	1.07	1.08	1.19	1.27	1.47	1.31	1.17	1.19	1.31	1.49
G(2,2)	1	123.20	1.68	3.69	1.85	1.46	1.20	1.10	1.09	1.12	1.18	1.60	1.34	1.20	1.19	1.24
G(3,1)	1	125.43	1.75	3.83	1.77	1.40	1.15	1.06	1.06	1.09	1.16	1.53	1.29	1.16	1.16	1.22
G(3,2)	1	134.06	1.27	2.27	2.14	1.66	1.30	1.13	1.05	1.05	1.03	1.82	1.48	1.23	1.14	1.11
U(.25,7)	1	118.74	1.58	3.47	1.97	1.55	1.26	1.14	1.13	1.16	1.21	1.70	1.41	1.25	1.23	1.26
h_{GLR}	6.3259	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
h_{EARS}	—	5.5702	—	—	—	—	—	—	—	—	—	—	—	—	—	—
$h_{Shewhart}$	—	—	3.5365	4.2454	—	—	—	—	—	—	—	—	—	—	—	—
h_{CUSUM}	—	—	—	—	21.8678	14.9425	9.8455	7.4639	6.3735	5.5281	4.8193	—	—	—	—	—
h_{EWMA}	—	—	—	—	—	—	—	—	—	—	—	2.7185	3.1101	3.7440	4.2905	4.8000

superior to the GLR chart when $\delta > 5$, but obviously inferior to the GLR chart when $\delta < 2$.

It is well known that CUSUM charts are generally efficient in detecting shifts with small value of δ, especially when δ is close to δ_1. This is also true when CUSUM charts are compared with the GLR chart. For example, the CUSUM chart with $\delta_1 = 1$ (column [7] in Table 9.1) is superior to the GLR chart at $\delta = 0.75, 1$, and 1.5 in terms of CED values, but performs not as well as the GLR chart elsewhere.

EWMA charts are also supposed to respond to small shifts quickly. Among the five EWMA chart settings in Table 9.1, only two of them (i.e., column [12] with $\alpha = 0.05$ and column [13] with $\alpha = 0.1$) are superior to the GLR chart in a narrow range of δ values according to CED values. From Table 9.1, we also see that CUSUM charts have better overall performance than EWMA charts based on our simulations.

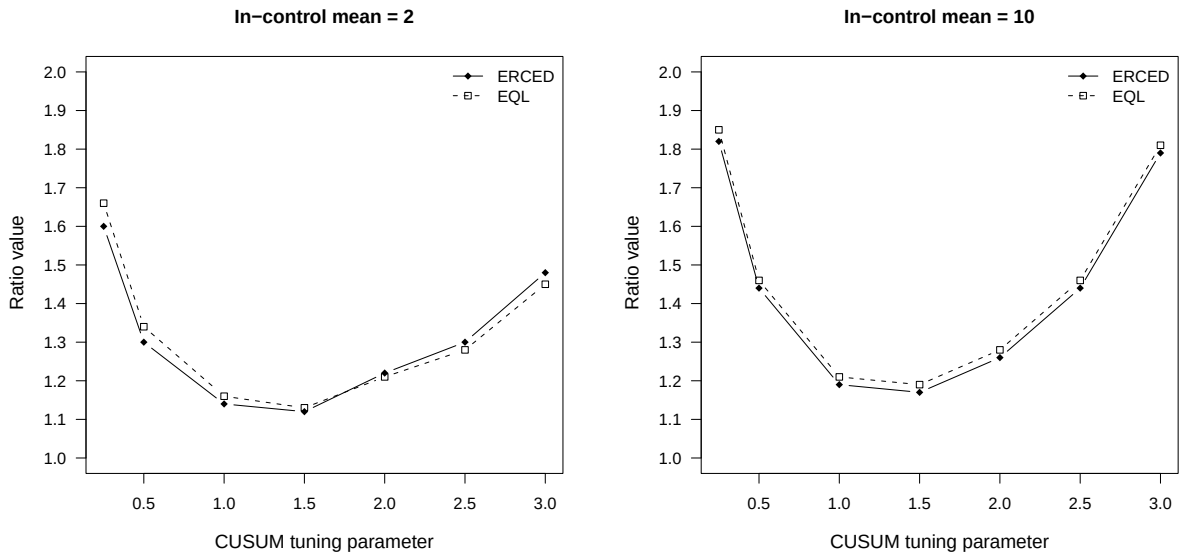
For GLR, CUSUM and EWMA charts in Table 9.1, their EQL values generally tend to increase with δ. However, if the shift size is as small as 0.25, the loss values may become larger, which results from the relatively large CED value for small shifts as indicated by Equation (6.19). In contrast, the EQL values of Shewhart charts tend to decrease with δ, which has been well studied and discussed by Reynolds *et al.* [86].

It is more intuitive, sometimes, to compare the ratios of the metrics of two methods than to compare their original metric values directly. EQL and ERCED ratios are also listed in Table (9.1), where ratios are defined by the values of alternative charts divided by their corresponding values of GLR charts. To evaluate the overall chart performance when δ varies, we assume the prior distribution $\pi(\delta)$ follows the gamma and uniform distributions. As discussed in Section 6.1, the gamma distribution is often used to describe the pattern of change in the mean in biosurveil-

lance. We apply the gamma distribution under seven different parameter settings, i.e., $G(1,1), G(1,2), G(1,3), G(2,1), G(2,2), G(3,1), G(3,2)$. The three settings with $a = 1$ are exponential distributions. The uniform distribution is used to represent the case in which all shifts in the interval [0.25, 7] are equally likely. The values of CED and $EQL(\delta)$ in Table 9.1 are based on the shift δ, while EQL ratio values and ERCED ratio values are based on additional simulations, in which the shifts are transformed Gaussian quadrature nodes as listed in Table 6.2. It can be seen that all EQL and ERCED ratios are greater than 1 across all gamma distributions and the uniform distribution considered, which indicates that the GLR chart has good overall performance compared with the alternative methods.

Table 9.1 also indicates the evaluation consistency between the EQL method and the ERCED method. For example, for the distribution $G(1,3)$, as the CUSUM design parameter δ_1 increases from 0.25 to 3, EQL ratios change from 1.66 to 1.45. We can see the corresponding ERCED ratios appear in the same pattern as the EQL ratios which are plotted in Figure 9.1(a). We also can see the difference between EQL ratios and ERCED ratios is very small for each δ_1. In fact, the consistency of EQL and ERCED results can also be found for $\lambda_0 = 10$ as illustrated in Figure 9.1(b), for which the numerical results are displayed later.

As shown in Equations (6.18) and (6.19), the EQL evaluation method involves the expectation calculation EQL(δ). For the sustained shift that we studied, this expectation involves the first moment of the time required to signal as shown in Equation (6.17), in this case, it is easy to calculate the expectation. However, if the mean shift follows a linear function, then the expectation will involve the third moment of the time required to signal [47], so the derivation for the EQL value will become more complicated. In contrast, computing ERCED values can avoid the computational complexity as shown in Equation (6.20), which leads to an ERCED evaluation method easy to use for different mean shift



(a) **(b)**

Figure 9.1: Comparison of EQL and ERCED evaluation methods when $\pi(\delta)$ is G(1, 3) and CUSUM δ_1 ranging from 0.25 to 3. (a) $\lambda_0 = 2$; (b) $\lambda_0 = 10$.

types. Furthermore, the evaluation performance of the ERCED method is similar to the EQL method. As discussed in Section 6.3, the EQL method has incorporated three effects involved in the process monitoring (i.e., loss, detection delay and shift distribution), while the ERCED method has considered two effects (i.e., detection delay and shift distribution), so if the process loss due to the out-of-control condition is not a concern, then the ERCED method is recommended, especially when the shift pattern is more complicated.

9.2 Comparison under different in-control means

Considering the discreteness of Poisson data, we changed $\lambda_0 = 2$ to $\lambda_0 = 10$ to investigate if the GLR chart performance changes with λ_0. Results for $\lambda_0 = 10$ are summarized in Table 9.2, in which simulation settings were similar to those settings for Table 9.1.

From Table 9.2, we can see that compared with the GLR chart, Shewhart charts perform well for large δ values and CUSUM charts perform well for those δ values which are close to the values of δ_1. Also, EWMA chart performance is not as good as CUSUM chart performance. EARS charts with extremely large CED values over the whole range of δ have the worst detection performance. Overall good performance of the GLR chart for $\lambda_0 = 10$ can be confirmed by the fact that all values of EQL and ERCED are greater than 1.

Figures 9.2 and 9.3 visually demonstrate some conclusions by combining the information of Tables 9.1 and 9.2. Figures 9.2(a) is a plot of the GLR CED value against δ for $\lambda_0 = 2$ and $\lambda_0 = 10$. It can be seen that the GLR chart performance for $\lambda_0 = 2$ and $\lambda_0 = 10$ is very similar in terms of CED values, so the CED value of the GLR Poisson chart seems to depend more on δ than on λ_0. This similarity between $\lambda_0 = 2$ and $\lambda_0 = 10$ results from use of the standardized shift in the mean. This happens to some exact with all detection methods. The EQL values versus δ are plotted in Figure 9.2(b), in which there is some difference in EQL values between $\lambda_0 = 2$ and $\lambda_0 = 10$ for large δ values, even though their CED values are very close each other. This results from the negative relationship between EQL and λ_0 as represented in Equation (6.19). Similar features discussed above can be found for the CUSUM with $\delta_1 = 1.5$ charts in Figures 9.2(c) and 9.2(d).

Tables 9.1 and 9.2 indicate that Shewhart and CUSUM charts can compete with the GLR chart for some mean shift sizes. To simplify the comparison between the GLR chart and the CUSUM chart, we used three typical tuning parameter settings for CUSUM charts; $\delta_1 = 0.5, 1.5$ and 2.5 represent the small, medium and large levels of δ_1, respectively. In order to make the CED comparisons more meaningful, the CED ratios, defined by the CED values of alternative charts divided by the corresponding CED values of the GLR chart for the same ATS_0 and δ, are used to evaluate the relative chart performance

Table 9.2: CED, EQL and ERCED values for charts with $\lambda_0 = 10$ for sustained shifts in the process mean

	GLR	EARS	Shewhart		CUSUM							EWMA				
δ_1 =	—	—	—	—	0.25	0.5	1	1.5	2	2.5	3	—	—	—	—	—
α =	—	—	—	—	—	—	—	—	—	—	—	0.05	0.1	0.2	0.3	0.4
δ	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
0.00	1502.34	1507.69	1411.18	3343.80	1500.71	1500.62	1522.77	1519.57	1522.59	1514.03	1855.91	1500.07	1500.70	1502.37	1501.88	1512.48
CED values																
0.25	116.63	1576.37	552.03	1217.27	80.16	102.53	173.44	247.49	311.28	381.15	516.64	117.73	162.46	240.23	308.45	364.06
0.50	37.37	1562.04	239.90	491.09	31.19	30.92	41.52	60.96	84.03	114.25	163.27	34.16	41.17	61.32	85.00	109.63
0.75	18.70	1524.78	116.62	226.36	18.92	16.89	18.00	22.94	30.22	41.61	59.86	18.28	18.83	24.04	32.18	41.28
1.00	11.50	1559.78	62.48	113.39	13.61	11.54	10.84	12.11	14.52	18.99	26.07	12.35	11.67	12.99	15.89	20.06
1.50	5.88	1509.79	21.40	36.09	8.68	6.98	5.83	5.68	6.00	6.88	8.37	7.50	6.52	6.22	6.64	7.45
2.00	3.67	1483.12	9.18	13.89	6.37	5.01	3.91	3.59	3.54	3.69	4.15	5.39	4.51	3.97	3.91	4.07
3.00	1.90	1318.36	2.68	3.65	4.14	3.16	2.33	1.99	1.86	1.77	1.82	3.46	2.79	2.29	2.07	1.97
4.00	1.20	1070.22	1.21	1.49	3.06	2.30	1.65	1.37	1.24	1.14	1.11	2.56	2.01	1.58	1.39	1.27
5.00	0.83	770.43	0.74	0.85	2.43	1.80	1.29	1.02	0.89	0.82	0.77	2.04	1.58	1.22	1.04	0.92
6.00	0.65	483.62	0.58	0.63	2.02	1.50	1.03	0.78	0.68	0.64	0.61	1.68	1.30	0.97	0.81	0.71
7.00	0.55	264.85	0.53	0.54	1.72	1.30	0.81	0.63	0.58	0.55	0.54	1.45	1.10	0.80	0.66	0.59
EQL(δ) values																
0.25	27.91	377.20	132.09	291.26	19.18	24.53	41.50	59.22	74.48	91.20	123.62	28.17	38.87	57.48	73.81	87.11
0.50	22.56	942.80	144.79	296.40	18.83	18.66	25.06	36.79	50.72	68.96	98.54	20.62	24.85	37.01	51.30	66.17
0.75	20.44	1666.30	127.45	247.37	20.68	18.46	19.67	25.07	33.03	45.47	65.42	19.98	20.58	26.27	35.17	45.11
1.00	19.63	2662.70	106.66	193.57	23.23	19.70	18.51	20.67	24.79	32.42	44.50	21.08	19.92	22.18	27.13	34.25
1.50	19.47	4998.40	70.85	119.48	28.74	23.11	19.30	18.81	19.86	22.78	27.71	24.83	21.59	20.59	21.98	24.66
2.00	19.87	8029.90	49.70	75.20	34.49	27.13	21.17	19.44	19.17	19.98	22.47	29.18	24.42	21.49	21.17	22.04
3.00	21.13	14661.90	29.81	40.59	46.04	35.14	25.91	22.13	20.69	19.69	20.24	38.48	31.03	25.47	23.02	21.91
4.00	22.59	20150.60	22.78	28.05	57.62	43.31	31.07	25.80	23.35	21.46	20.90	48.20	37.85	29.75	26.17	23.91
5.00	23.68	21984.60	21.12	24.26	69.34	51.36	36.81	29.11	25.40	23.40	21.97	58.21	45.09	34.81	29.68	26.25
6.00	26.16	19462.10	23.34	25.35	81.29	60.36	41.45	31.39	27.37	25.76	24.55	67.61	52.32	39.04	32.60	28.57
7.00	29.67	14288.60	28.59	29.13	92.79	70.14	43.70	33.99	31.29	29.67	29.13	78.23	59.35	43.16	35.61	31.83
EQL ratios																
G(1,1)	1	204.95	4.71	9.04	1.25	1.08	1.11	1.33	1.63	2.05	2.78	1.18	1.18	1.38	1.70	2.05
G(1,2)	1	372.48	3.59	6.54	1.64	1.33	1.16	1.22	1.37	1.62	2.08	1.46	1.30	1.31	1.47	1.67
G(1,3)	1	453.45	3.06	5.41	1.85	1.46	1.21	1.19	1.28	1.46	1.81	1.61	1.39	1.31	1.39	1.52
G(2,1)	1	389.91	3.41	6.02	1.67	1.34	1.13	1.14	1.25	1.45	1.83	1.46	1.28	1.24	1.36	1.52
G(2,2)	1	593.61	2.14	3.42	2.20	1.68	1.27	1.13	1.10	1.15	1.29	1.86	1.52	1.28	1.23	1.25
G(3,1)	1	569.76	2.26	3.59	2.08	1.60	1.22	1.09	1.07	1.13	1.27	1.76	1.44	1.24	1.20	1.22
G(3,2)	1	710.55	1.41	1.96	2.56	1.92	1.38	1.14	1.04	1.01	1.03	2.14	1.69	1.34	1.19	1.12
U(.25,7)	1	612.97	1.99	3.17	2.35	1.79	1.33	1.16	1.14	1.18	1.32	1.99	1.61	1.34	1.27	1.27
ERCED ratios																
G(1,1)	1	211.55	4.67	8.88	1.27	1.09	1.09	1.30	1.57	1.98	2.67	1.19	1.17	1.35	1.65	1.99
G(1,2)	1	367.98	3.61	6.53	1.63	1.32	1.15	1.20	1.34	1.59	2.03	1.44	1.29	1.29	1.44	1.64
G(1,3)	1	443.73	3.12	5.49	1.82	1.44	1.19	1.17	1.26	1.44	1.79	1.58	1.36	1.29	1.37	1.51
G(2,1)	1	383.46	3.44	6.03	1.66	1.33	1.12	1.13	1.23	1.43	1.80	1.45	1.26	1.23	1.34	1.50
G(2,2)	1	578.57	2.23	3.58	2.14	1.64	1.25	1.12	1.10	1.16	1.31	1.82	1.48	1.27	1.23	1.25
G(3,1)	1	554.48	2.33	3.73	2.04	1.57	1.20	1.08	1.07	1.13	1.28	1.73	1.42	1.22	1.19	1.22
G(3,2)	1	702.26	1.48	2.09	2.50	1.88	1.36	1.13	1.04	1.01	1.04	2.10	1.66	1.32	1.19	1.12
U(.25,7)	1	600.67	2.09	3.36	2.28	1.75	1.31	1.15	1.13	1.19	1.33	1.94	1.57	1.33	1.27	1.27
h_{GLR}	6.3954	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
h_{EARS}	—	6.6265	—	—	—	—	—	—	—	—	—	—	—	—	—	—
$h_{Shewhart}$	—	—	3.7046	3.7948	—	—	—	—	—	—	—	—	—	—	—	—
h_{CUSUM}	—	—	—	—	47.5361	31.4063	19.4569	14.4728	11.3807	9.7155	8.5615	—	—	—	—	—
h_{EWMA}	—	—	—	—	—	—	—	—	—	—	—	11.5551	12.3723	13.6321	14.7042	15.6802

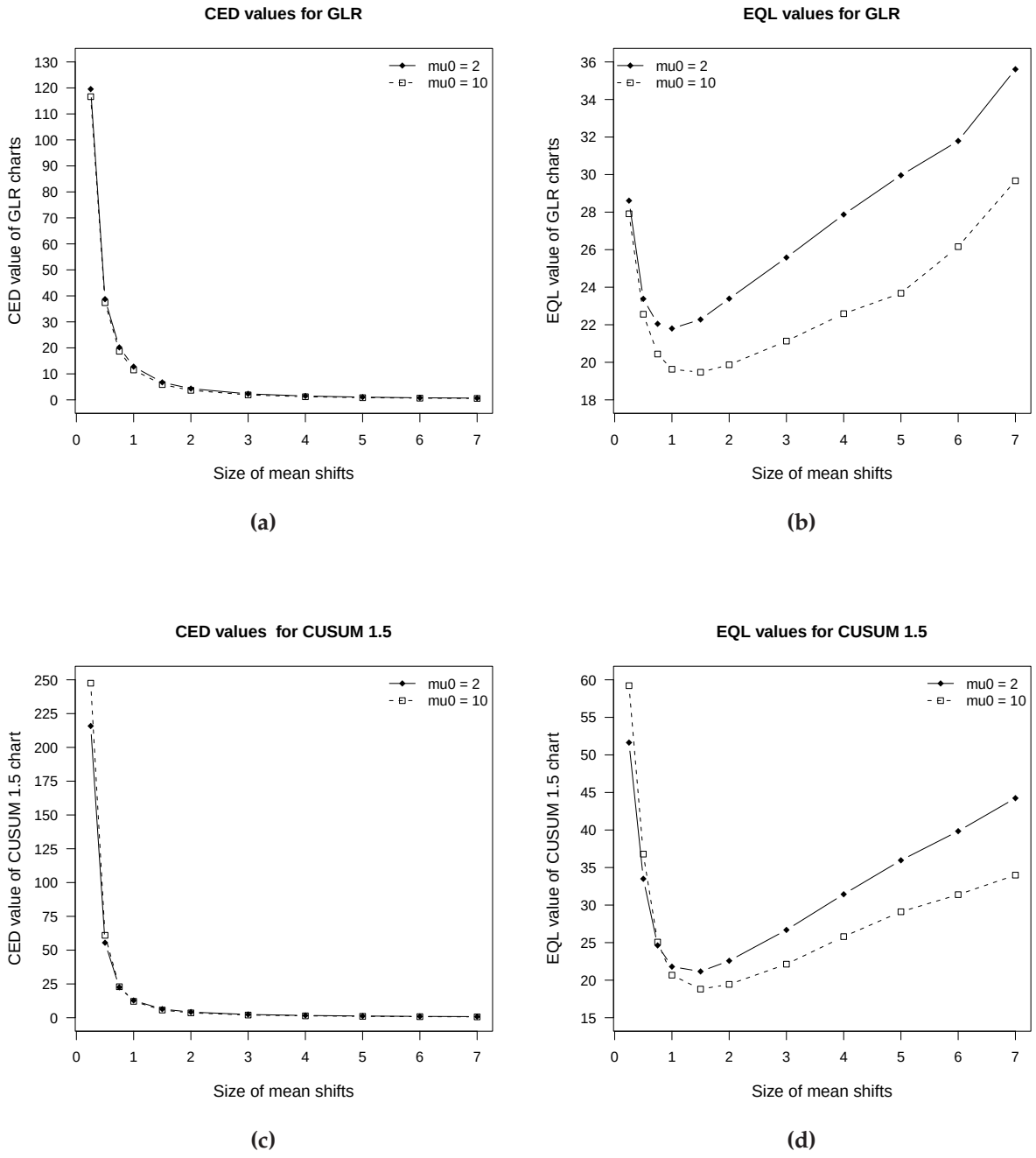


Figure 9.2: Comparison of CED and EQL values between $\lambda_0 = 2$ and $\lambda_0 = 10$. (a) GLR CED values; (b) GLR EQL values; (c) CED values of CUSUM with $\delta_1 = 1.5$; (d) EQL values of CUSUM with $\delta_1 = 1.5$.

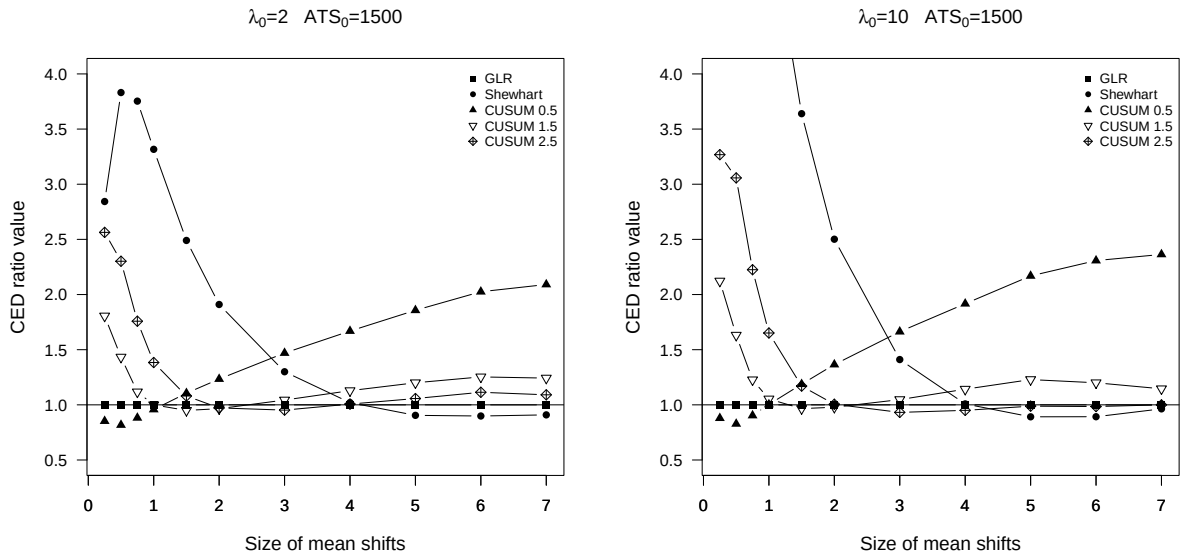


Figure 9.3: CED performance comparison of GLR, Shewhart and CUSUM charts when $\text{ATS}_0 = 1500$, where CED ratios were the values of the GLR chart's CED divided by alternative methods' CED. (a) $\lambda_0 = 2$; (b) $\lambda_0 = 10$.

according to CED.

Figures 9.3(a) and 9.3(b) plot the CED ratios of GLR, Shewhart and three CUSUM charts against a range of mean shifts for $\lambda_0 = 2$ and $\lambda_0 = 10$, respectively. As shown in these two figures, the Shewhart method performs well for large values of δ while it performs not as well for small values of δ. CUSUM charts perform well when δ_1 values are close to δ. Plots for Shewhart and CUSUM charts cross the reference line (i.e., CED ratio $= 1$), which means that no single chart can be superior to the GLR chart over the entire range of mean shifts. The very similar trend patterns shown in Figures 9.3(a) and 9.3(b) imply that the chart performance seems not depend on the Poisson in-control mean.

9.3 Comparison using AMOC

In previous sections, the performance comparisons between different charts were based on $ATS_0 = 1500$. While such a high ATS_0 value increases the performance for in-control conditions, it decreases the sensitivity of the disease outbreak detection in biosurveillance practice. Thus, a wide range of smaller ATS_0 values were also explored for $\lambda_0 = 2$ to study the chart performance and to examine the consistency of chart performance for different values of ATS_0.

Tables 9.3 and 9.4 give CED values for combinations of ATS_0 and δ for one-sided GLR, Shewhart, CUSUM, EARS and EWMA charts when $\lambda_0 = 2$. In these two tables, the control limit of each chart that was needed to obtain the specified ATS_0 was first estimated based on the linear relationship as described in Section 8.2, and then determined more accurately by simulation. Control limits for control charts in these tables, as well as for other control charts, are summarized in Appendix C, which can be used as a reference table for various control charts based on their desired in-control performance.

From Tables 9.3 and 9.4, as a whole, the EARS method has very poor performance when compared with other methods, and EWMA charts seems inferior to CUSUM charts over the entire range of ATS_0 values considered. As a result, the EARS and EWMA methods will not be compared with the GLR and CUSUM methods in later study. Results in Table 9.3 are summarized by varying the shift sizes δ in Figure 9.4, and by varying ATS_0 values as in Figure 9.5.

Figure 9.4 is a set of four AMOC curves for Poisson data with $\lambda_0 = 2$, where we chose four typical mean shifts (i.e., $\delta = 0.5, 1.5, 3$ and 6) to represent a wide range of δ, and selected three typical CUSUM tuning parameter settings (i.e., $\delta_1 = 0.5, 1.5$ and 2.5) to represent a range of δ_1. As expected and shown in these four graphs, CUSUM charts

Table 9.3: CED values for combinations of δ and ATS_0 for one-sided GLR, Shewhart, CUSUM charts with $\lambda_0 = 2$

			Size of mean shifts (δ)										
	ATS_0	FAR	0.25	0.50	0.75	1.00	1.50	2.00	3.00	4.00	5.00	6.00	7.00
GLR	100	0.0100	18.68	11.21	7.13	4.99	2.96	2.04	1.21	0.85	0.67	0.58	0.54
	150	0.0067	24.56	14.02	8.77	5.91	3.37	2.25	1.28	0.88	0.69	0.59	0.54
	200	0.0050	29.18	16.12	9.89	6.66	3.83	2.56	1.48	1.02	0.78	0.65	0.58
	300	0.0033	51.11	21.57	12.19	7.95	4.35	2.87	1.62	1.09	0.81	0.66	0.59
	400	0.0025	62.36	24.79	13.88	8.99	4.87	3.18	1.73	1.13	0.82	0.67	0.59
	600	0.0017	76.17	28.42	15.47	9.98	5.37	3.52	1.98	1.34	0.98	0.76	0.65
	800	0.0012	84.34	30.77	16.56	10.65	5.73	3.72	2.04	1.36	0.98	0.77	0.65
	1000	0.0010	92.70	33.05	17.67	11.28	6.03	3.90	2.12	1.39	1.00	0.77	0.65
	1200	0.0008	108.85	36.22	18.94	11.94	6.24	3.99	2.14	1.40	1.00	0.78	0.65
	1500	0.0007	119.55	38.74	20.18	12.77	6.73	4.32	2.30	1.48	1.05	0.79	0.66
Shewhart	100	0.0100	94.93	47.63	27.23	16.70	7.69	4.22	1.82	1.05	0.74	0.61	0.55
	150	0.0067	94.93	47.63	27.23	16.70	7.69	4.22	1.82	1.05	0.74	0.61	0.55
	200	0.0050	94.93	47.63	27.23	16.70	7.69	4.22	1.82	1.05	0.74	0.61	0.55
	300	0.0033	94.93	47.63	27.23	16.70	7.69	4.22	1.82	1.05	0.74	0.61	0.55
	400	0.0025	94.93	47.63	27.23	16.70	7.69	4.22	1.82	1.05	0.74	0.61	0.55
	600	0.0017	94.93	47.63	27.23	16.70	7.69	4.22	1.82	1.05	0.74	0.61	0.55
	800	0.0012	339.92	148.42	75.74	42.35	16.76	8.25	2.99	1.51	0.95	0.71	0.60
	1000	0.0010	339.92	148.42	75.74	42.35	16.76	8.25	2.99	1.51	0.95	0.71	0.60
	1200	0.0008	339.92	148.42	75.74	42.35	16.76	8.25	2.99	1.51	0.95	0.71	0.60
	1500	0.0007	339.92	148.42	75.74	42.35	16.76	8.25	2.99	1.51	0.95	0.71	0.60
CUSUM 0.5	100	0.0100	24.70	11.74	7.18	5.05	3.12	2.25	1.46	1.08	0.86	0.72	0.63
	150	0.0067	32.32	14.24	8.58	5.97	3.65	2.63	1.69	1.25	0.99	0.82	0.70
	200	0.0050	38.36	16.23	9.64	6.72	4.11	2.96	1.90	1.40	1.11	0.91	0.78
	300	0.0033	48.66	19.22	11.21	7.74	4.71	3.37	2.15	1.58	1.25	1.03	0.87
	400	0.0025	56.79	21.38	12.39	8.55	5.21	3.74	2.38	1.75	1.39	1.14	0.97
	600	0.0017	68.68	24.41	14.01	9.67	5.84	4.17	2.64	1.94	1.54	1.27	1.08
	800	0.0012	77.90	26.50	15.13	10.42	6.31	4.53	2.88	2.11	1.67	1.37	1.18
	1000	0.0010	85.95	28.43	16.07	11.08	6.71	4.81	3.05	2.24	1.77	1.46	1.24
	1200	0.0008	93.25	29.80	16.85	11.61	7.05	5.06	3.21	2.36	1.85	1.52	1.30
	1500	0.0007	101.92	31.66	17.80	12.23	7.44	5.33	3.38	2.47	1.95	1.60	1.38
CUSUM 1.5	100	0.0100	33.31	15.76	9.14	5.91	3.25	2.17	1.32	0.94	0.74	0.63	0.57
	150	0.0067	45.13	19.15	10.67	6.78	3.70	2.49	1.47	1.02	0.78	0.65	0.58
	200	0.0050	58.55	23.45	12.47	7.72	4.06	2.67	1.58	1.12	0.87	0.71	0.62
	300	0.0033	73.65	27.15	13.80	8.49	4.47	2.99	1.76	1.22	0.92	0.74	0.63
	400	0.0025	88.88	31.19	15.15	9.23	4.76	3.13	1.79	1.24	0.93	0.74	0.64
	600	0.0017	116.24	37.63	17.22	10.33	5.25	3.49	2.05	1.43	1.08	0.86	0.71
	800	0.0012	141.90	42.61	18.83	11.08	5.61	3.67	2.09	1.45	1.09	0.86	0.72
	1000	0.0010	167.22	48.31	20.60	11.86	5.85	3.80	2.22	1.59	1.22	0.97	0.80
	1200	0.0008	195.15	52.53	21.81	12.41	6.19	4.06	2.37	1.66	1.26	0.99	0.81
	1500	0.0007	215.82	55.50	22.54	12.77	6.39	4.17	2.40	1.67	1.26	0.99	0.82
CUSUM 2.5	100	0.0100	41.17	19.72	11.16	7.05	3.65	2.30	1.28	0.87	0.69	0.59	0.54
	150	0.0067	56.42	25.71	14.02	8.53	4.18	2.59	1.39	0.92	0.71	0.60	0.54
	200	0.0050	74.08	31.14	16.01	9.54	4.56	2.84	1.55	1.05	0.80	0.66	0.58
	300	0.0033	91.09	37.78	18.44	10.81	5.00	3.03	1.60	1.07	0.81	0.66	0.59
	400	0.0025	114.80	43.78	20.45	11.71	5.31	3.24	1.71	1.13	0.83	0.67	0.59
	600	0.0017	150.81	54.46	23.97	13.23	5.82	3.51	1.87	1.27	0.95	0.75	0.64
	800	0.0012	193.99	65.39	28.13	14.94	6.40	3.78	1.95	1.30	0.96	0.76	0.64
	1000	0.0010	193.99	65.39	28.13	14.94	6.40	3.78	1.95	1.30	0.96	0.76	0.64
	1200	0.0008	264.00	79.89	32.25	16.48	6.93	4.06	2.10	1.39	1.00	0.78	0.65
	1500	0.0007	306.47	89.17	35.48	17.67	7.27	4.20	2.19	1.49	1.11	0.88	0.72

Table 9.4: CED values for combinations of ATS_0 and δ for one-sided EARS and EWMA charts with $\lambda_0 = 2$

			Shift size (δ)										
	ATS_0	FAR	0.25	0.50	0.75	1.00	1.50	2.00	3.00	4.00	5.00	6.00	7.00
EARS	100	0.0100	90.98	78.77	70.82	62.43	48.32	36.39	18.65	8.02	3.65	1.49	0.88
	150	0.0067	129.65	113.90	103.68	91.79	73.75	57.70	31.15	15.05	6.93	2.73	1.39
	200	0.0050	167.98	148.92	133.94	118.53	96.76	78.37	43.13	22.19	10.17	4.06	1.79
	300	0.0033	253.33	224.64	203.94	184.58	159.87	129.18	80.81	47.85	22.05	9.75	4.26
	400	0.0025	310.31	272.12	246.08	222.66	191.67	158.90	102.53	60.50	28.08	13.38	5.69
	600	0.0017	469.35	403.43	357.40	324.02	269.61	223.24	146.70	90.10	45.97	22.04	10.30
	800	0.0012	636.79	538.57	472.68	423.15	354.25	299.51	203.68	130.01	69.85	35.84	15.95
	1000	0.0010	789.40	666.65	578.42	528.50	438.11	370.15	251.83	166.63	89.61	46.04	20.41
	1200	0.0008	917.93	778.11	674.66	614.11	514.67	449.98	318.30	210.75	122.62	63.21	30.55
	1500	0.0007	1111.92	933.11	795.50	714.37	600.09	517.36	371.26	246.19	151.13	82.28	42.20
EWMA 0.05	100	0.0100	26.98	13.52	8.77	6.45	4.21	3.16	2.12	1.61	1.29	1.08	0.92
	150	0.0067	34.40	16.15	10.23	7.43	4.78	3.56	2.37	1.79	1.43	1.19	1.02
	200	0.0050	40.58	18.10	11.28	8.16	5.21	3.85	2.55	1.91	1.53	1.27	1.09
	300	0.0033	51.68	21.28	12.89	9.20	5.80	4.25	2.80	2.09	1.68	1.38	1.18
	400	0.0025	60.30	23.53	14.01	9.89	6.21	4.54	2.97	2.22	1.76	1.46	1.25
	600	0.0017	74.67	26.96	15.64	10.87	6.77	4.93	3.20	2.39	1.89	1.56	1.34
	800	0.0012	86.55	29.60	16.73	11.55	7.15	5.20	3.35	2.50	1.98	1.64	1.40
	1000	0.0010	96.77	31.73	17.64	12.09	7.44	5.39	3.48	2.58	2.05	1.69	1.44
	1200	0.0008	105.88	33.33	18.37	12.54	7.67	5.55	3.57	2.65	2.10	1.74	1.48
	1500	0.0007	118.07	35.60	19.31	13.10	7.96	5.74	3.69	2.73	2.16	1.79	1.52
EWMA 0.1	100	0.0100	28.99	14.00	8.73	6.21	3.92	2.86	1.86	1.39	1.10	0.91	0.78
	150	0.0067	38.14	16.93	10.17	7.11	4.40	3.18	2.05	1.51	1.20	0.98	0.83
	200	0.0050	45.75	19.13	11.24	7.78	4.71	3.41	2.18	1.61	1.26	1.04	0.88
	300	0.0033	58.70	22.64	12.82	8.70	5.21	3.72	2.37	1.73	1.36	1.12	0.94
	400	0.0025	70.96	25.57	14.00	9.39	5.56	3.94	2.49	1.82	1.43	1.17	0.99
	600	0.0017	91.06	30.30	15.82	10.38	6.04	4.25	2.66	1.94	1.53	1.24	1.05
	800	0.0012	110.09	34.07	17.18	11.07	6.37	4.46	2.78	2.02	1.59	1.29	1.09
	1000	0.0010	127.13	37.19	18.31	11.66	6.64	4.63	2.88	2.09	1.64	1.34	1.13
	1200	0.0008	141.39	39.86	19.23	12.12	6.85	4.76	2.96	2.14	1.68	1.37	1.15
	1500	0.0007	157.08	42.89	20.15	12.60	7.06	4.90	3.02	2.19	1.72	1.40	1.18
EWMA 0.2	100	0.0100	32.45	15.21	9.09	6.19	3.68	2.56	1.60	1.16	0.90	0.74	0.64
	150	0.0067	43.92	19.05	10.87	7.17	4.13	2.86	1.75	1.26	0.98	0.79	0.68
	200	0.0050	53.39	22.14	12.18	7.92	4.46	3.07	1.86	1.34	1.03	0.83	0.71
	300	0.0033	71.03	27.16	14.13	8.94	4.91	3.35	1.99	1.43	1.10	0.88	0.74
	400	0.0025	88.66	31.87	15.76	9.87	5.26	3.55	2.10	1.50	1.15	0.92	0.77
	600	0.0017	118.49	39.55	18.52	11.11	5.78	3.82	2.25	1.59	1.22	0.98	0.81
	800	0.0012	144.78	45.41	20.58	12.01	6.14	4.02	2.34	1.65	1.27	1.01	0.84
	1000	0.0010	170.77	51.44	22.37	12.83	6.45	4.19	2.43	1.71	1.31	1.04	0.87
	1200	0.0008	197.59	56.87	24.14	13.57	6.69	4.34	2.50	1.75	1.34	1.07	0.89
	1500	0.0007	233.35	63.75	26.26	14.47	7.00	4.50	2.58	1.80	1.38	1.10	0.91

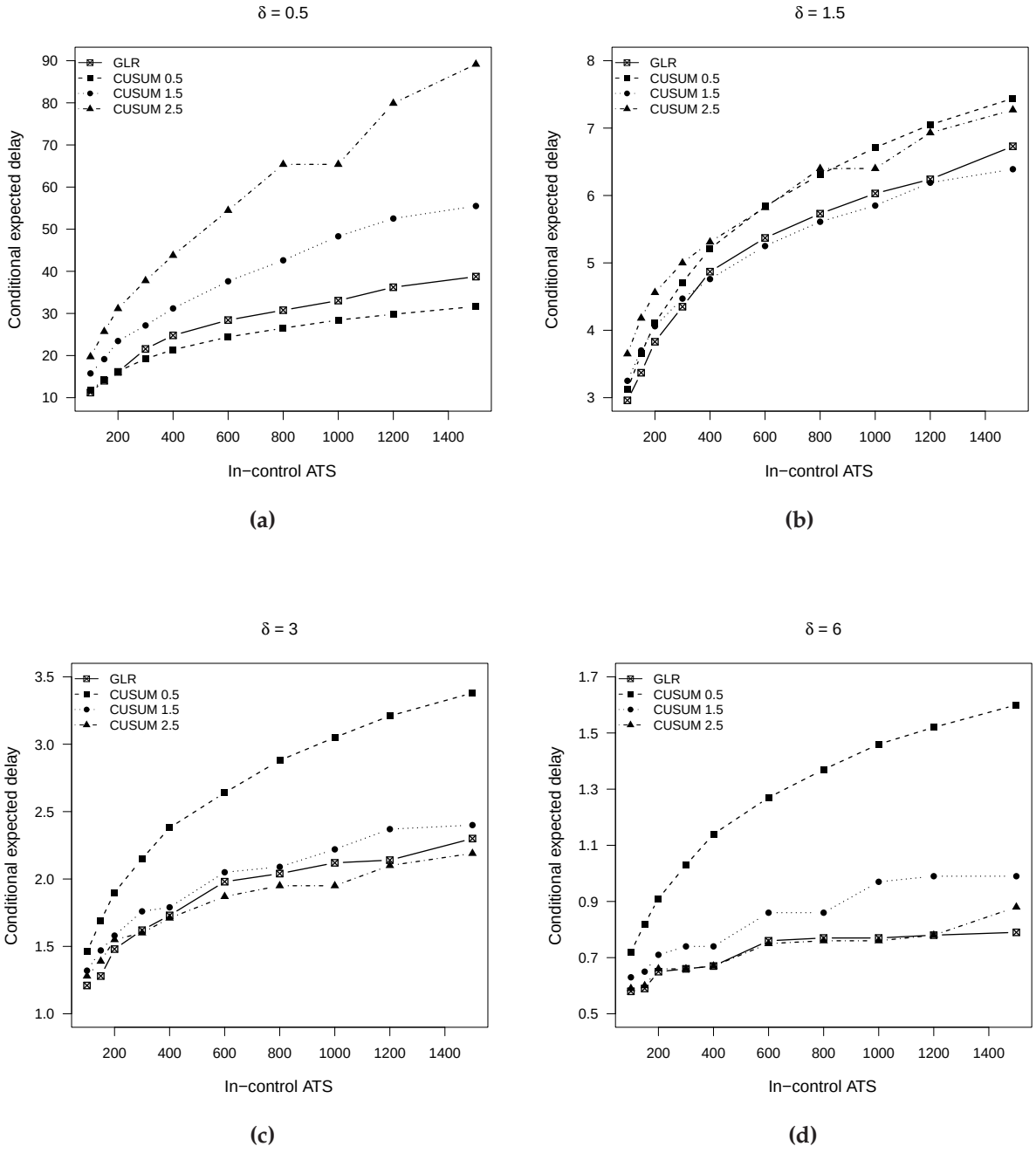


Figure 9.4: A set of four AMOC curves for Poisson data with $\lambda_0 = 2$. (a) $\delta = 0.5$; (b) $\delta = 1.5$; (c) $\delta = 3$; (d) $\delta = 6$.

perform much better when δ_1 is close to the shift size δ than when δ_1 is far away from δ. In general, CUSUM charts designed with δ_1 values close to δ perform better than the GLR chart at the higher ATS_0 values, but the CUSUM charts perform only as well as or even worse than the GLR chart at relatively smaller ATS_0 values. For example, in Figure 9.4(b) where $\delta = 1.5$, the CUSUM chart with $\delta_1 = 1.5$ is superior to the GLR chart when $ATS_0 \geq 400$ and inferior to the GLR chart when $ATS_0 < 400$.

The performance of GLR, Shewhart and CUSUM charts for $ATS_0 = 1500$ in terms of their CED values has been discussed in Section 9.2. With the results for the multiple ATS_0 values available, these discussion can be extended to a wider set of circumstances.

The CED performance of charts of interest listed in Table 9.3 are visually plotted in Figure 9.5, in which four choices of ATS_0 represent the practical levels of ATS_0 for biosurveillance. From Figure 9.5(a), it can be seen that CED ratio values of the Shewhart and CUSUM charts are greater than 1, indicating the GLR chart appears superior to the Shewhart and CUSUM charts across the entire range of δ when $ATS_0 = 100$. CED values of the Shewhart and CUSUM charts become lower than the GLR chart in some ranges of δ, but no individual chart is superior to the GLR chart over the whole δ range, as illustrated in Figures 9.5(b), (c) and (d). Combining the information given in Figure 9.5, we can see that, in general, the Shewhart chart performs well when $\delta > 4$; the CUSUM chart with $\delta_1 = 0.5$ can outperform the GLR chart when $\delta < 1$; the CUSUM chart with $\delta_1 = 1.5$ and 2.5 performs well when $\delta \in [1, 4]$.

Based on Figures 9.3, 9.4 and 9.5, we may conclude that the GLR chart has overall good performance over the varying mean shifts and ATS_0 values, even though for some values of δ, alternative charts may be a bit better than the GLR chart with respect to the CED value.

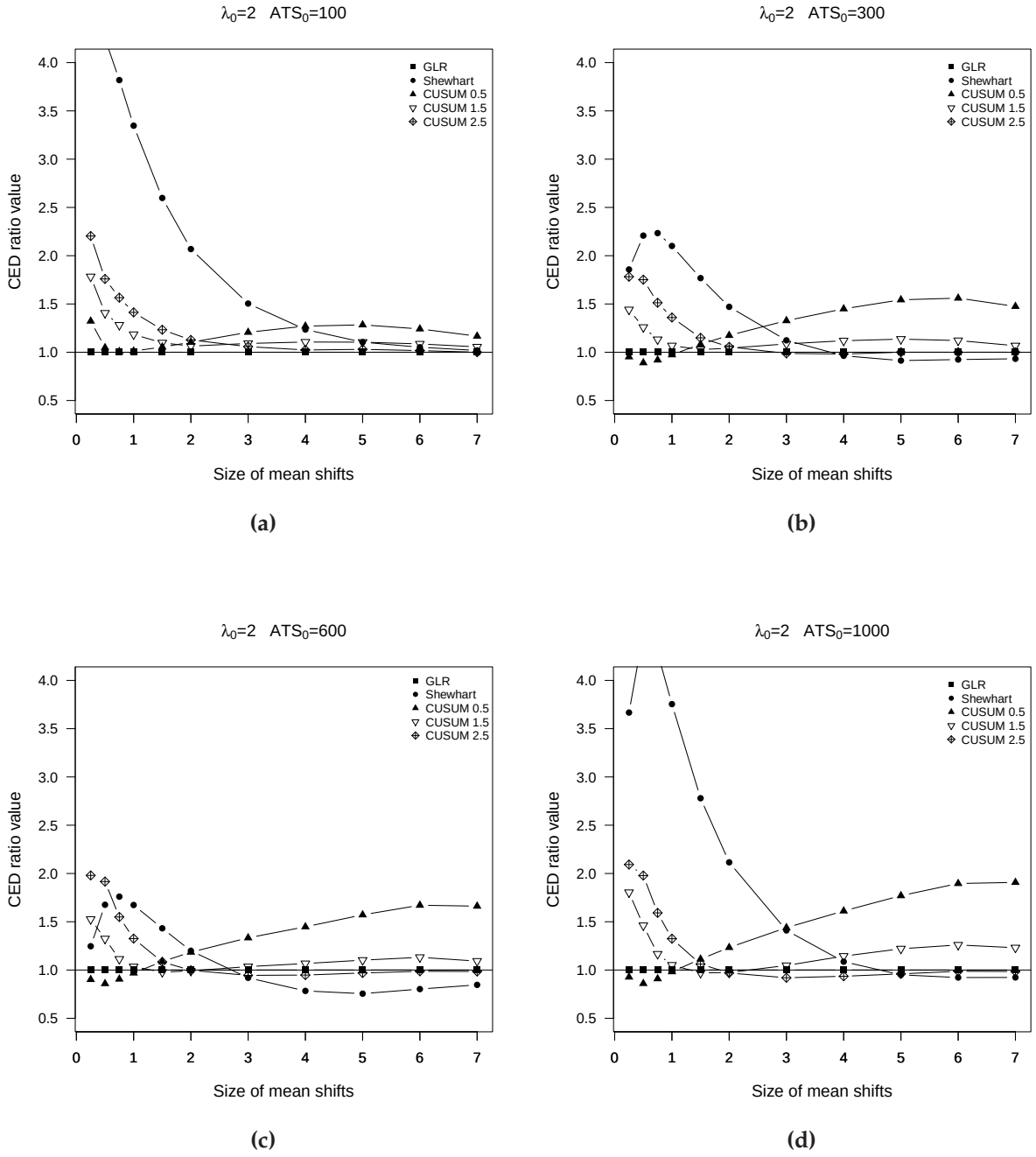


Figure 9.5: CED performance comparison of GLR, Shewhart and CUSUM charts for Poisson data with $\lambda_0 = 2$, where CED ratios are the CED values of alternative methods divided by those of the GLR chart. (a) ATS_0=100; (b) ATS_0=300; (c) ATS_0=600; (d) ATS_0=1000.

Chapter 10

Conclusions and future work

10.1 Conclusions

In the biosurveillance context, it is essential to detect increases in disease rates as soon as possible once an increase occurs in the population. Recognizing the increase sooner may allow officials to react more promptly to prevent further spread of the disease. It follows that any improvement in average time to signal performance of a monitoring approach should be pursued. The GLR chart proposed in this part can be applied in detecting a wide range of increases in rates when the magnitude of the increase is unknown.

In Part II, we derived surveillance statistics for both two-sided and one-sided GLR charts for Poisson counts, and suggested a practical window length to ease the computational work associated with GLR statistics. We also demonstrated the calculation methods for determining control limits, and some example GLR charts. Performance comparisons of GLR charts to other charts were then made under a variety of circumstances, such as various magnitudes of the in-control mean, shift sizes, chart parameters, and ATS_0 values.

Because GLR statistics without the use of a window are maximized over all past samples, when the sample size becomes large, the computational work will consequently be

time-consuming. To speed the calculations, we proposed that GLR charts with a window of length 400 that can provide nearly optimal performance compared to larger window lengths. The general relationship between the window size m and the in-control ATS_0 and shift size δ that I we concluded can be confirmed theoretically by Lai [95]. This paper extended Lorden's [78] asymptotic properties of the GLR chart without the window and implied that the GLR chart with a moving window is also asymptotically optimal if the window size $m = O(\ln \gamma / I_{\min})$, where γ is the ATS_0 and $I_{\min}$ is the Kullback-Leibler information number corresponding to the minimum shift of interest. Consequently, we can reach the conclusion that a larger value of m is needed if the desired value of ATS_0 is large or the target shift size is small, and a smaller value of m is sufficient when the value of ATS_0 is small or the expected shift size is large.

In some applications, however, the GLR chart with $m = 400$ may be not suitable. For example, when changing the sustained shift to a transient shift in which the out-of-control duration is short, a smaller length window is better than a larger one. See Reynolds *et al.* [47] for more details.

We modified the two-sided GLR statistics to one-sided GLR statistics with a sign function, which functionally leads to two-sided, upward-sided and downward-sided charts sharing one uniform expression. Although the chart performance of the alternative modification which transforms two-sided statistics to one-sided statistics by using a reflecting barrier is the same as that of the modification with the sign function, the recommended surveillance statistics for GLR charts are the sign function applied statistics, i.e., $R^s_{m,k}$.

To get the control limits for specified performance efficiently, we suggest to take full advantage of the linear relationship between the control limits and ATS_0 on the log scale. Furthermore, to facilitate GLR chart users, control limits for $\lambda_0 = 2$ were simulated, and

summarized in Table C, which covers a practical biosurveillance range of ATS_0 values.

In Part II, we conclusively demonstrated that overall the GLR chart outperforms the EARS, Shewhart, CUSUM and EWMA approaches under the steady-state average time to signal framework in a wide variety of circumstances. The key results of the comparisons are summarized as follows:

1. The surveillance statistics of the Shewhart chart and GLR chart with $m = 1$ are equivalent in terms of their CDF. The theoretical proof is given in Appendix B. Consequently, the surveillance performance of these two charts are same.

2. The EQL and ERCED evaluation approaches provide similar results, but the EQL method is simpler than the ERCED method. Due to computational complexity (e.g., the calculation of expectation) involved in EQL, the ERCED method will be a good choice.

3. The performance of GLR charts mainly depends on the mean shift size rather than the in-control mean size. If CED values of two GLR charts are similar, the GLR chart with larger in-control mean will have less expected loss than the GLR chart with smaller mean.

4. The GLR chart relative performance is stable over a wide range of ATS_0 values. In particular, the GLR chart is superior to alternative charts investigated when the ATS_0 value is around 100.

5. The EARS method has the worse performance among all methods. In the research framework of the steady-state average time to signal, EWMA charts are inferior to CUSUM charts over a range of ATS_0 and δ values. Note that the EARS method is at a large disadvantage since the in-control parameter is not assumed to be known.

6. Compared with all alternative charts considered in this study, the proposed GLR chart appears to have overall good performance under a variety of circumstances.

10.2 Future work

All results shown thus far are based on comparisons of the GLR chart with individual traditional control charts. As stated earlier, the GLR chart outperforms alternative charts over a wide range of shifts, but we also showed the Shewhart chart performs well for large shifts while CUSUM charts are superior to GLR charts when CUSUM tuning parameters are close to the actual shift sizes. Thus, it would be of future research interest to compare the GLR chart with combinations of the Shewhart chart and CUSUM chart(s), in which each individual chart is designed to obtain its optimal performance. Fortunately, Figures 9.3 and 9.5 have provided some information for designing these kinds of combinations. In fact, the Shewhart chart is a special case of the CUSUM chart, and the GLR chart can be viewed as an infinite combination of CUSUM charts, see Reynolds *et al.* [47] for more discussion. A new study would be to use combinations of CUSUM charts to approximate the GLR chart performance.

Other than combinations of CUSUM charts, there are other choices to detect a range of mean shifts, such as ACUSUM-C charts, and WCUSUM-C charts. The ACUSUM-C chart was proposed by Jiang [35] and developed based on the adaptive CUSUM charts, i.e., ACUSUM [34]. The ACUSUM-C chart adaptively estimates the process mean by using the EWMA statistic to improve the efficiency of detecting an unknown δ, and it has been used in recent papers [85, 96]. More specifically, by the ACUSUM-C approach, the

estimator of λ_1 at sample k can be expressed as

$$\hat{\lambda}_{1,k} = \hat{\lambda}_{1,k-1} + \phi(e_k),$$

where $e_k = x_k - \hat{\lambda}_{1,k-1}$ is the prediction error, and $\phi(\cdot)$ is a monotonic function, which can be represented as

$$\phi(e) = \begin{cases} e + (1-\theta)\gamma\sqrt{\lambda_0} & e < -\gamma\sqrt{\lambda_0} \\ \theta e & |e| \leq \gamma\sqrt{\lambda_0} \\ e - (1-\theta)\gamma\sqrt{\lambda_0} & e > \gamma\sqrt{\lambda_0}. \end{cases}$$

The initial value of $\hat{\lambda}_{1,0}$ is usually set at zero but other positive values could be chosen if fast initial response of the ACUSUM chart is desired [31].

Since detecting the upward shift in λ is the main objective in our current study, to prevent the estimate of λ from falling below a specified minimum value of interest, we can replace $\hat{\lambda}_{1,k}$ by $\hat{\lambda}^+_{1,k} = \max\{\lambda^+_{\min}, \hat{\lambda}_{1,k}\}$, where $\hat{\lambda}^+_{1,k}$ is another design parameter of the ACUSUM-C chart. Now that the estimator of the process mean at sample k is $\hat{\lambda}^+_{1,k}$, the surveillance statistic for the ACUSUM-C chart becomes

$$AC^+_k = \max\{0, AC^+_{k-1} + x_k - r^+_k\},$$

where

$$r^+_k = \frac{\hat{\lambda}^+_{1,k} - \lambda_0}{\ln\hat{\lambda}^+_{1,k} - \ln\lambda_0}.$$

Note that $\lambda^+_{\min}$ can be represented in a standardized form, $\delta_{\min} = (\lambda^+_{\min} - \lambda_0)/\sqrt{\lambda_0}$, so there are three parameters involved in the ACUSUM-C chart, i.e., θ, γ and $\delta_{\min}$.

The WCUSUM-C chart was proposed by Shu *et al.* [36] for patterned mean shifts.

This chart can be treated as a modified version of the ACUSUM-C chart. While the ACUSUM-C chart keeps updating r_k^+ as $\hat{\lambda}_1$ is estimated, the WCUSUM-C chart fixes the reference value of CUSUM and updates a weight function for $\hat{\delta}_k$. Compared with the ACUSUM-C chart, the WCUSUM-C chart needs one more parameter, i.e., the fixed reference value, which needs to be predetermined optimally to produce good chart performance [36]. Because of the simplicity compared with the WCUSUM-C chart, we can compare the GLR chart with ACUSUM-C charts under various settings of ACUSUM-C parameters in future work.

To evaluate the GLR chart performance in depth, there may be many other choices besides comparing the GLR chart with ACUSUM-C charts, e.g., changes in the type of mean shifts. In our study thus far, we assumed that a special cause resulted in the mean shift from λ_0 to an unknown value that persisted until detected by a control chart and action was taken to remove the special cause. This is generally referred to as a sustained shift. In some applications, the shift in λ may be in a gradual shift pattern instead of a sudden jump. For example, if a linear shift in λ that starts at time τ^* between τ and $\tau+1$, then the process mean can be represented as

$$\lambda(t) = \lambda_0 + \beta(t-\tau^*)\sqrt{\lambda_0},$$

where $t \geq \tau^*$, and β is the standardized rate of the shift. This kind of linear shift is generally referred to as a drift in the mean. Evaluating chart performance on drifts and various levels of β will be another interest of our future study.

In the context of biosurveillnce, however, the types of disease outbreak signatures that have been used usually span a short number of days, not because the disease is expected to disappear, but rather because early detection is to be evaluated. One type of

signature shape is a constant number of cases added for a few consecutive days [97, 98]. As a result, limiting the duration of out-of-control condition to a few time periods will be an additional modification of the sustained shift, for which we may need to change the chart evaluation metric from CED, EQL and ERCED to other appropriate criteria, such as the probability of successful detection (PSD) of mean changes. Furthermore, GLR charts with window length of $m = 400$ may be not suitable for the temporary shifts [42, 47], so we will adjust the window length to optimize the GLR chart, and compare its performance with alternatives.

The performance evaluation of GLR charts was based on the one-sided control charts thus far. In that the two-sided GLR chart statistics are available, we can examine the performance of two-sided GLR charts in future studies.

The Poisson distribution is often used to model health-care related data, such as visits to an emergency department room. However, in some cases, the data are frequently characterized by an excessive number of zeros, for example, the observed perinatal deaths in a county [99]. Zero-inflated Poisson (ZIP) models proposed by Lambert [100] provide a way to model this type of situation. This next part is a study of the GLR algorithm for ZIP data.

Part III

GLR Charts for ZIP Data

Chapter 11

Introduction

Although the Poisson distribution is widely used for the analysis of count data, its application has been restricted by the lack of flexibility resulting from having only one parameter. Examples include the variance exceeding the mean, the variance being less than the mean, and the presence of an excessive number of zero counts. In the medical context, a possible explanation for the excess of zero counts might be due to the fact when all patients are cured after the treatment, no realization of the symptom being monitored will occur [101]. Failure to account for extra zeros may lead to underprediction of the true frequency of zeros, overprediction of the true frequency of other small values, and underprediction the true frequency of large counts [102]. To take into account extra zeros, a zero-inflated Poisson (ZIP) distribution considers a mixture model, in which extra zeros are modeled by a Bernoulli distribution and other counts are assumed to follow a Poisson distribution. There are two parameters involved in a ZIP model, the probability for the Bernoulli distribution and the mean of the Poisson distribution.

The ZIP model has attracted more attention in various scientific fields since the publication by Lambert [100], who applied the ZIP regression model in a manufacturing defect case study. Dietz [103] proposed a zero-modified Poisson model, which is a generalization of the Poisson model, zero-inflated Poisson model, and zero-deflated Poisson model.

This work also provided an approach to estimate the Poisson mean of the zero-modified Poisson model by truncating the Poisson data. Since such estimation is derived under the assumption of sufficient sample sizes, one should use caution when applying this proposed method. Relaxing the assumption of data independence, Agarwal *et al.* [104] considered the ZIP model in the context of correlated data in space, and Lee *et al.* [105] investigated a hierarchical model for correlated ZIP data. ZIP regression models have been used in diverse research fields, such as use of recreational facilities [102], road safety [106], patent applications [107], and sexual behavior [108].

In the medical context, Farewell *et al.* [109] used a ZIP model with no regressor to analyze the counts of premature ventricular contractions (PVC). Fong *et al.* [110] analyzed the same PVC data set using an alternative algorithm (EM) and produced the similar results. Other applications of the ZIP model include fetal movement [111], fetal growth and postnatal somatic growth [112], dental epidemiology [113], and disease patterns [114].

It is known that the negative binomial distribution can be viewed as a Poisson distribution under which the Poisson mean follows a gamma distribution. As a natural generalization of the ZIP model, a zero-inflated negative binomial (ZINB) model provides more flexibility in modeling nonzero counts. Applications and more discussion on ZINB models can be found in the recent literature [114, 115, 116, 117].

Compared with the Poisson model, the ZIP model incorporates inflation of the rate of zeros through an additional parameter. A ZIP model is thus applied only when Poisson models are inappropriate. Tests for ZIP models against Poisson models and ZINB models against ZIP models have been proposed in the statistical literature; see for example [118, 119, 120, 121]. Xie *et al.* [122] summarized the results of some of these tests including the score test, the likelihood ratio test, the Chi-square test, the Cochran test, and the Rao-Chakravarti test. Performance of these tests has been evaluated when two parameters are

at relatively large magnitudes; see Xie *et al.*'s [122] paper for details.

In contrast to extensive applications and discussion of ZIP regression models, statistical process control charts for ZIP data have not been as throughly investigated. Xie *et al.* [122, 123] proposed a ZIP-Shewhart chart to detect shifts in both parameters of a ZIP process. Recently, He *et al.* [124] developed a set of CUSUM charts to monitor the individual parameters and both ZIP parameters simultaneously (i.e., ZIP-CUSUM charts). Under the general CUSUM framework, this work used p-CUSUM charts and λ-CUSUM charts to monitor the Bernoulli parameter and the Poisson mean, respectively. To detect shifts in both parameters, the authors used either a combination of p-CUSUM chart and λ-CUSUM chart, or a single chart called the t-CUSUM chart. Surveillance performance of such CUSUM charts was compared and compared with the ZIP-Shewhart chart. They pointed out that the ZIP-Shewhart chart is insensitive to shifts in the Bernoulli parameter and smaller shifts in the Poisson mean, and also concluded that the combination method is much better than the single CUSUM chart when one parameter increases as the other decreases.

It is known that the best performance of a CUSUM chart can be achieved only when its tuning parameter corresponds to the real shift size, which requires practitioners to have some prior knowledge of the process. It has been generally recommended that the CUSUM reference value corresponds to the smallest shift size one is interested in detecting quickly. The chart statistics of the combination method proposed by He *et al.* consist of two tuning parameters, one for the Bernoulli parameter and the other for the Poisson mean. Mis-specification of any tuning parameter may lead to a loss of charts' power. Moreover, specifying these two unknown tuning parameters simultaneously requires an in-depth knowledge of the process, which can be challenging or even impossible in practice. For this reason, a control chart for monitoring two process parameters should detect

a wide range of shifts with little or no knowledge of the shift sizes that will occur. As demonstrated in Part II, the GLR control chart may be a good candidate for such a purpose.

Unlike the CUSUM method requiring pre-specified tuning parameters, the GLR algorithm uses the built-in maximum likelihood estimate (MLE) to estimate the process parameters of interest, followed by calculation of likelihood ratios. These likelihood ratios are then maximized using a changepoint framework, producing the GLR chart statistics. As a consequence, the GLR algorithm is often referred to as a double maximization scheme. A GLR chart can also be viewed as a combination of an infinite number of CUSUM charts, leading to the detection of a wide range of shifts in process parameters of interest. It has been shown that the GLR algorithm is conceptually simple but computationally complicated; see Part II and the paper by Reynolds *et al.* [47] for more details.

The original specific goals on this research of the GLR algorithm for ZIP processes can be summarized as follows:

1. Investigate the MLE performance, and make some modification or improvements if needed;
2. Develop the ZIP-GLR chart for various magnitudes of the in-control Bernoulli parameter and the in-control Poisson mean, and investigate the performance of GLR charts for varying levels of these parameters;
3. Compare the performance between the GLR chart and other charts;
4. Demonstrate how to create a GLR chart and how to interpret the results;
5. Provide guidelines on designing GLR charts in practice;
6. Provide reference tables of control limits for varying in-control performance;

7. Apply the GLR chart on sparse data in biosurveillance;

8. Develop an R package to facilitate use of the GLR chart, and publish it in the Comprehensive R Archive Network (CRAN).

Because the computational work involved in the GLR algorithm is extensive and time is limited, the study of GLR charts for ZIP data in the dissertation is the basis for future GLR chart developments. Specifically, we accomplished the estimation of the ZIP process parameters, performance investigation of the parameter estimation, and derivation and discussion of the chart statistics.

Chapter 12

Derivation of GLR control charts

The Poisson distribution is often used to model health-care related data, such as visits to an emergency department. When data are characterized by an excessive number of zeros, for example, the observed perinatal deaths in a county [99], zero-inflated Poisson (ZIP) models discussed by Lambert [100] can be applied.

In this chapter, we first derive MLEs for a set of GLR charts for ZIP data, and explore their features and performance using numerical and graphical approaches. We then derive the GLR statistics, followed by changing two-sided statistics to one-sided statistics with a sign function. We give discussion and suggestions about GLR charts throughout this chapter.

12.1 ZIP data features

The ZIP model can be viewed as a mixture model of a Bernoulli distribution with parameter p and a Poisson distribution with mean λ, in which some zero and all nonzero counts are results from outcomes of the Poisson distribution, so the value of p can be viewed as

the index of zero inflation. A ZIP model can be formulated as

$$f(x;p,\lambda) = \begin{cases} 1-p+pe^{-\lambda} & x=0 \\ p\dfrac{\lambda^x e^{-\lambda}}{x!} & x>0, \end{cases} \tag{12.1}$$

where values of x are integers and $0 \leq p \leq 1$. If $p = 0$, all counts will be zeros; if $p = 1$, the ZIP model will reduce to an ordinary Poisson distribution with mean λ. The mean and variance of a ZIP distribution are the following:

$$E(X) = p\lambda, \qquad Var(X) = p\lambda + \lambda^2 p(1-p). \tag{12.2}$$

We can see that zero counts of a ZIP distribution can be categorized into two sources: zeros originating from the Bernoulli distribution which are called structural zeros, and zeros produced by the Poisson distribution which are called sampling zeros [125]. The proportion of sampling zeros increases as the value of λ decreases. Understanding features and properties of ZIP data in different parameter settings will be valuable for parameter estimation, especially when sample sizes are small.

To gain insight into the ZIP data structure, an investigation of such data was performed. It can be seen that the sparseness of ZIP data is evident in Figure 12.1(a) where $p = 0.1$ and $\lambda = 2$. Even in cases where both p and λ are relatively large, an excessive number of zeros can still be found in Figure 12.1(b). Figures 12.1(c) and (d) demonstrate how the probability of zero counts changes with the value of p and λ, respectively. For a given value of p, the difference in the density between $\lambda = 2$ and $\lambda = 5$ is apparent as shown in Figure 12.1(c). However, if λ is relatively large, such as $\lambda \geq 5$, the density mainly depends on the value of p as indicated by the two overlapping curves in Figure 12.1(c). This fact can be confirmed by the near equality of the probability of zero counts

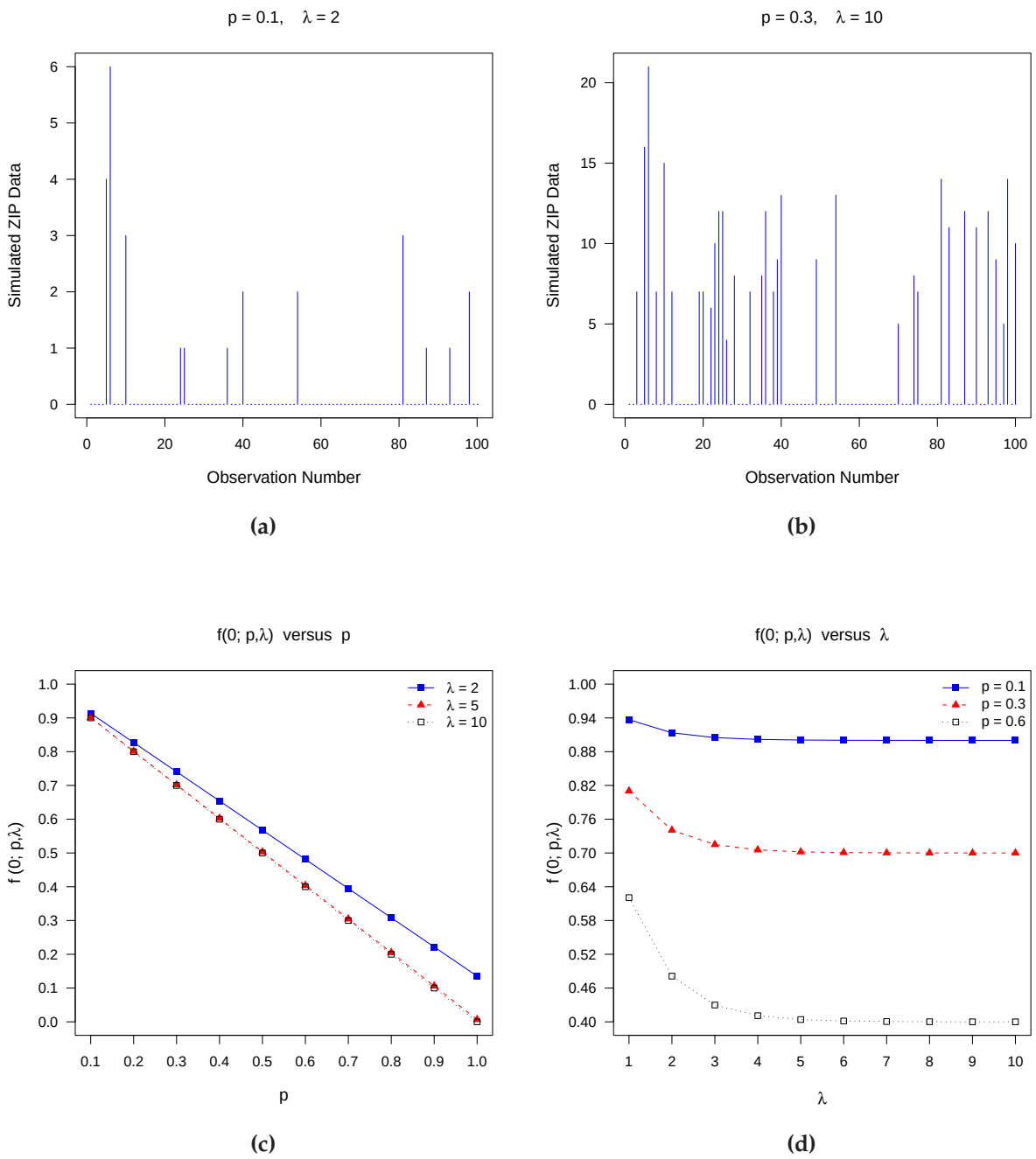


Figure 12.1: Demonstration of ZIP data features. (a) simulated ZIP data when $p = 0.1$, $\lambda = 2$; (b) simulated ZIP data when $p = 0.3$, $\lambda = 10$; (c) the probability of zero counts as a function of p; (d) the probability of zero counts as a function of λ.

when $\lambda \geq 5$ as shown in Figure 12.1(d). A close examination of Figure 12.1(d) also reveals that the number of structural zeros considerably dominates the number of sampling zeros in ZIP data when $\lambda \geq 5$.

In the following sections, MLEs for ZIP parameters will be derived, followed by the development of chart statistics and the investigation of their properties.

12.2 Likelihood function of ZIP data

Suppose a sequential sample with size n, $\{X_1, X_2, \cdots, X_n\}$, is obtained from a ZIP process. Let $l = \max\{X_1, X_2, \cdots, X_n\}$, then this sample can be described by its count frequency. Let $n_0, n_1, \cdots, n_l$ be the frequencies of counts 0, 1, 2, $\cdots$, l, respectively. Then the sample size n can be expressed as

$$n = n_0 + \sum_{j=1}^{l} n_j, \tag{12.3}$$

and the number of positive counts will be

$$\sum_{j=1}^{l} n_j = n - n_0. \tag{12.4}$$

Consequently, the mean of the positive counts, $\overline{x}^+$, is given by

$$\overline{x}^+ = \frac{\sum_{j=1}^{l} j \cdot n_j}{n - n_0}. \tag{12.5}$$

The log likelihood function for the sample $\{X_1, X_2, \cdots, X_n\}$ can then be expressed as

$$\begin{aligned} \ln L &= n_0 \ln f(0; p, \lambda) + \sum_{j=1}^{l} n_j \ln f(j; p, \lambda) \\ &= n_0 \ln(1 - p + pe^{-\lambda}) + \sum_{j=1}^{l} n_j (\ln p - \lambda) + \sum_{j=1}^{l} j \cdot n_j \ln\lambda - \sum_{j=1}^{l} n_j \ln(j!). \end{aligned} \tag{12.6}$$

Using Equations (12.4) and (12.5), Equation (12.6) can be reduced to

$$\ln L = n_0 \ln(1 - p + pe^{-\lambda}) + (n - n_0)(\ln p - \lambda + \overline{x}^{+} \ln\lambda) - \sum_{j=1}^{l} n_j \ln(j!). \tag{12.7}$$

The GLR chart is a plot of its statistics against the time sequence. As demonstrated in Part II, the derivation of chart statistics is based on the double maximization scheme. Therefore, the accuracy of the MLE can have a significant effect on the GLR chart performance. Moreover, small sizes are not avoidable due to the changepoint framework, which may result in some performance issues for the MLEs.

12.3 Derivation of p-GLR chart statistics

In this section, we derive the MLE for the p-GLR chart, which is designed to detect shifts in the Bernoulli parameter p, when the Poisson mean λ is assumed to remain a known constant. The accuracy of the MLE and properties of the p-GLR chart statistics will then be investigated.

12.3.1 MLE for p-GLR charts

If we assume the value of λ remains unchanged over a ZIP process and only consider shifts in p , then based on Equation (12.7) the score equation for the p-GLR chart is

$$\frac{\partial \ln L}{\partial p} = \frac{n_0(e^{-\lambda} - 1)}{1 - p + pe^{-\lambda}} + \frac{n - n_0}{p}, \tag{12.8}$$

and the resulting MLE for p is given by

$$\hat{p} = \frac{n - n_0}{n} \cdot \frac{1}{1 - e^{-\lambda}}. \tag{12.9}$$

We can see Equation (12.9) will be undefined if $\lambda = 0$, but this situation does not occur in practice. This equation indicates that the MLE for p is determined by the proportion of positive counts when the value of λ is a constant. However, it is possible that $\hat{p} > 1$ in Equation (12.9), which requires the use of an upper bound.

Figure 12.2 is a plot of values of $\hat{p}$ against the number of zero counts in a sample with size 20 for three levels of λ. It shows that for all choices of λ, $\hat{p} > 1$ when $n_0 = 0$. This inequality also holds in some cases where both λ values and sample sizes are small. In fact, such cases are possible for the GLR algorithm due to use of the built-in changepoint framework, which makes the sample size automatically range from 1 to n. To prevent values of $\hat{p}$ beyond the possible range, we can put an upper bound on $\hat{p}$ as follows

$$\hat{p}_B = \min(1, \hat{p}) = \min\left(1, \quad \frac{n - n_0}{n} \cdot \frac{1}{1 - e^{-\lambda}}\right). \tag{12.10}$$

The bounded MLE of p, i.e., $\hat{p}_B$, will be used for the development of p-GLR charts, and, for the sake of simplicity, we will still use the notation $\hat{p}$ instead of $\hat{p}_B$.

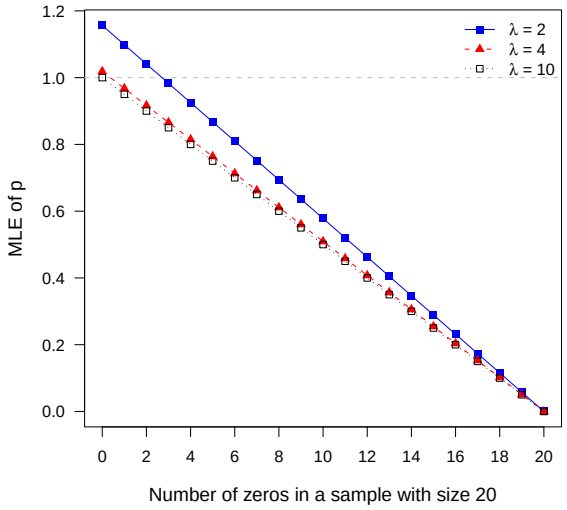


Figure 12.2: Relationship between $\hat{p}$ and n_0 in p-GLR charts.

MLEs play an important role in the GLR algorithm. Poor performance of the MLE, such as excessive variation, will result in unstable or unreliable chart performance. To explore the effect of small sample sizes on $\hat{p}$, we performed a set of simulations under various conditions, including varying values of p and λ. As demonstrated in later sections, the computational work involved in GLR charts for ZIP can be extensive especially when some iterative algorithms are applied, so time used in future simulation studies is a concern. In this sense, we also varied the number of simulations to look for a smaller simulation number, based on which the MLE performance is relatively close to that based on a greater number of simulations.

Table 12.1 reports simulation settings and the summarized results for $\lambda = 2$. Since the effect of small sample sizes is our primary concern, we selectively list results for sample sizes ranging from 1 to 30 in this table, but more information is displayed in Figure 12.3. To consider different magnitudes of p, values in this table are the ratios, $\hat{p}/p$, where the value of $\hat{p}$ is the bounded MLE for p.

Table 12.1: Effect of small sample sizes on the MLE for p in p-GLR charts, where $\lambda = 2$. The values in this table are of $\hat{p}/p$

	N.Simulation $= 1,000$				N.Simulation $= 5,000$				N.Simulation $= 10,000$				N.Simulation $= 50,000$			
Sample Size	$p = 0.1$	0.2	0.3	0.6	0.1	0.2	0.3	0.6	0.1	0.2	0.3	0.6	0.1	0.2	0.3	0.6
1	0.910	0.915	0.820	0.888	0.900	0.876	0.862	0.867	0.885	0.894	0.875	0.864	0.865	0.872	0.866	0.865
2	0.949	0.983	0.920	0.935	0.973	0.989	0.955	0.922	0.985	1.004	0.963	0.925	0.989	0.981	0.965	0.929
3	0.945	1.002	0.996	0.961	0.976	1.006	0.994	0.957	0.992	1.014	0.993	0.961	0.996	0.999	0.991	0.964
4	0.958	1.029	0.988	0.986	0.986	1.010	1.005	0.977	0.997	1.016	1.003	0.977	1.000	1.003	0.998	0.980
5	0.997	1.026	0.994	0.987	1.007	1.008	1.005	0.984	1.014	1.014	1.003	0.982	1.004	1.004	0.997	0.990
6	1.000	1.002	0.994	0.984	1.009	1.001	1.001	0.991	1.014	1.008	0.997	0.989	1.001	1.004	0.997	0.995
7	0.998	1.009	0.989	0.992	1.006	1.003	0.999	0.996	1.014	1.008	0.997	0.994	0.999	1.005	0.996	0.998
8	0.992	1.022	0.984	0.998	1.003	1.007	0.999	0.995	1.015	1.012	0.998	0.994	1.000	1.006	0.996	0.998
9	0.982	1.018	0.989	1.003	1.007	1.008	0.998	0.998	1.016	1.012	0.999	0.996	0.999	1.007	0.996	0.998
10	1.002	1.012	0.983	1.001	1.010	1.008	0.995	0.999	1.013	1.011	0.997	0.996	1.000	1.004	0.996	0.999
11	1.008	1.006	0.991	1.003	1.010	1.008	0.996	0.999	1.014	1.012	0.996	0.997	1.003	1.005	0.996	1.000
12	0.994	1.003	1.005	1.001	1.008	1.007	1.005	0.998	1.013	1.012	1.005	0.996	1.002	1.004	1.001	1.000
13	0.994	1.007	1.006	1.001	1.003	1.007	1.005	0.997	1.009	1.010	1.006	0.996	1.001	1.004	1.001	0.999
14	0.996	1.002	1.011	1.003	1.004	1.008	1.008	0.999	1.006	1.011	1.007	0.996	1.001	1.003	1.002	1.000
15	1.005	1.010	1.002	1.001	1.006	1.010	1.007	0.998	1.008	1.012	1.008	0.996	1.000	1.002	1.002	1.000
16	1.000	1.004	1.001	1.001	1.004	1.007	1.006	0.998	1.006	1.008	1.007	0.997	1.001	1.001	1.002	1.000
17	1.005	1.001	1.000	1.002	1.003	1.008	1.006	0.998	1.005	1.008	1.007	0.997	1.001	1.001	1.002	1.000
18	1.013	0.996	0.999	1.002	1.006	1.005	1.007	0.998	1.007	1.007	1.009	0.997	1.003	1.000	1.002	1.000
19	1.011	1.000	1.005	1.003	1.009	1.006	1.008	0.999	1.009	1.006	1.009	0.998	1.004	1.000	1.002	1.000
20	1.009	0.993	1.008	1.004	1.011	1.004	1.009	0.999	1.012	1.005	1.009	0.998	1.004	1.000	1.002	1.000
21	1.008	1.000	1.007	1.003	1.013	1.006	1.009	0.999	1.013	1.006	1.008	0.998	1.004	1.000	1.002	1.000
22	1.004	1.007	1.008	1.003	1.012	1.011	1.008	0.999	1.013	1.011	1.008	0.997	1.003	1.002	1.002	1.000
23	1.004	1.004	1.009	1.003	1.013	1.010	1.008	0.998	1.013	1.011	1.008	0.997	1.002	1.003	1.002	1.000
24	1.011	1.006	1.007	1.005	1.014	1.011	1.008	0.999	1.014	1.011	1.008	0.998	1.002	1.004	1.002	1.000
25	1.009	0.999	1.002	1.004	1.014	1.010	1.007	0.999	1.013	1.011	1.008	0.998	1.001	1.004	1.002	1.000
26	1.007	0.998	0.999	1.003	1.012	1.008	1.005	0.998	1.010	1.011	1.007	0.997	1.001	1.003	1.002	1.000
27	1.007	0.996	1.001	1.003	1.014	1.009	1.005	0.999	1.010	1.012	1.007	0.998	1.001	1.003	1.002	1.000
28	1.005	0.993	0.999	1.004	1.011	1.009	1.005	0.999	1.009	1.011	1.007	0.998	1.000	1.003	1.002	1.000
29	1.010	0.997	1.000	1.005	1.013	1.010	1.006	0.999	1.009	1.012	1.008	0.998	1.001	1.003	1.003	1.000
30	1.005	0.997	1.002	1.003	1.012	1.010	1.006	0.999	1.009	1.011	1.007	0.998	1.001	1.003	1.003	1.000

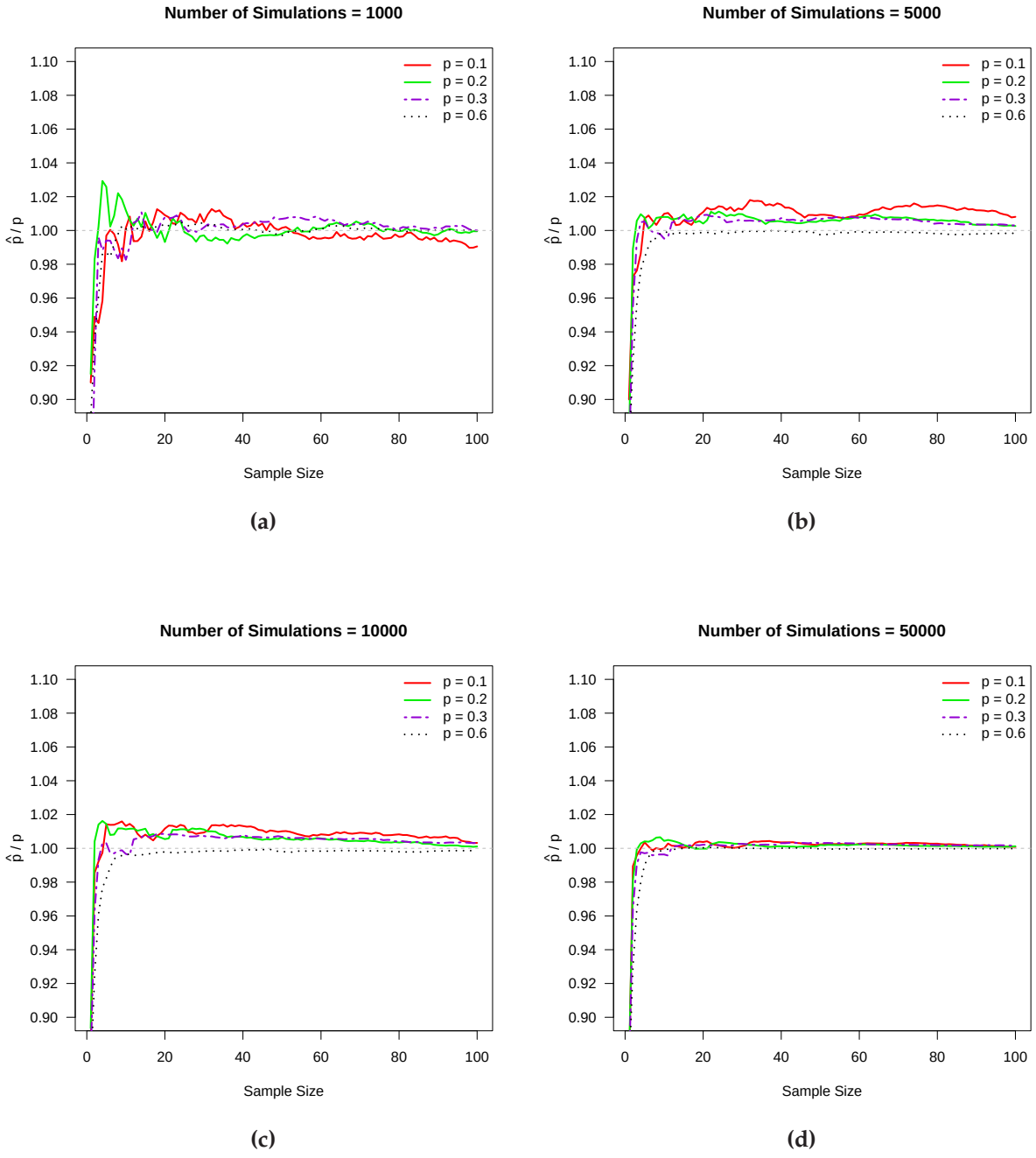


Figure 12.3: Effect of small sample sizes on the MLE for p in p-GLR charts, where $\lambda = 2$. (a) number of simulations $= 1,000$; (b) number of simulations $= 5,000$; (c) number of simulations $= 10,000$; (d) number of simulations $= 50,000$.

Figure 12.3 indicates that for most choices of p, we have $\hat{p}/p < 1$ when sample sizes are less than 5, implying a systematic underestimation of p if the value of λ is small. It can be seen that values of $\hat{p}/p$ approach 1 as sample sizes become larger. In particular, these values are almost always in the range $[0.98, 1.02]$ when sample sizes are greater than 20, and the MLEs yield good performance when sample sizes are close to 100. Moreover, the MLE performs better as the value of p increases.

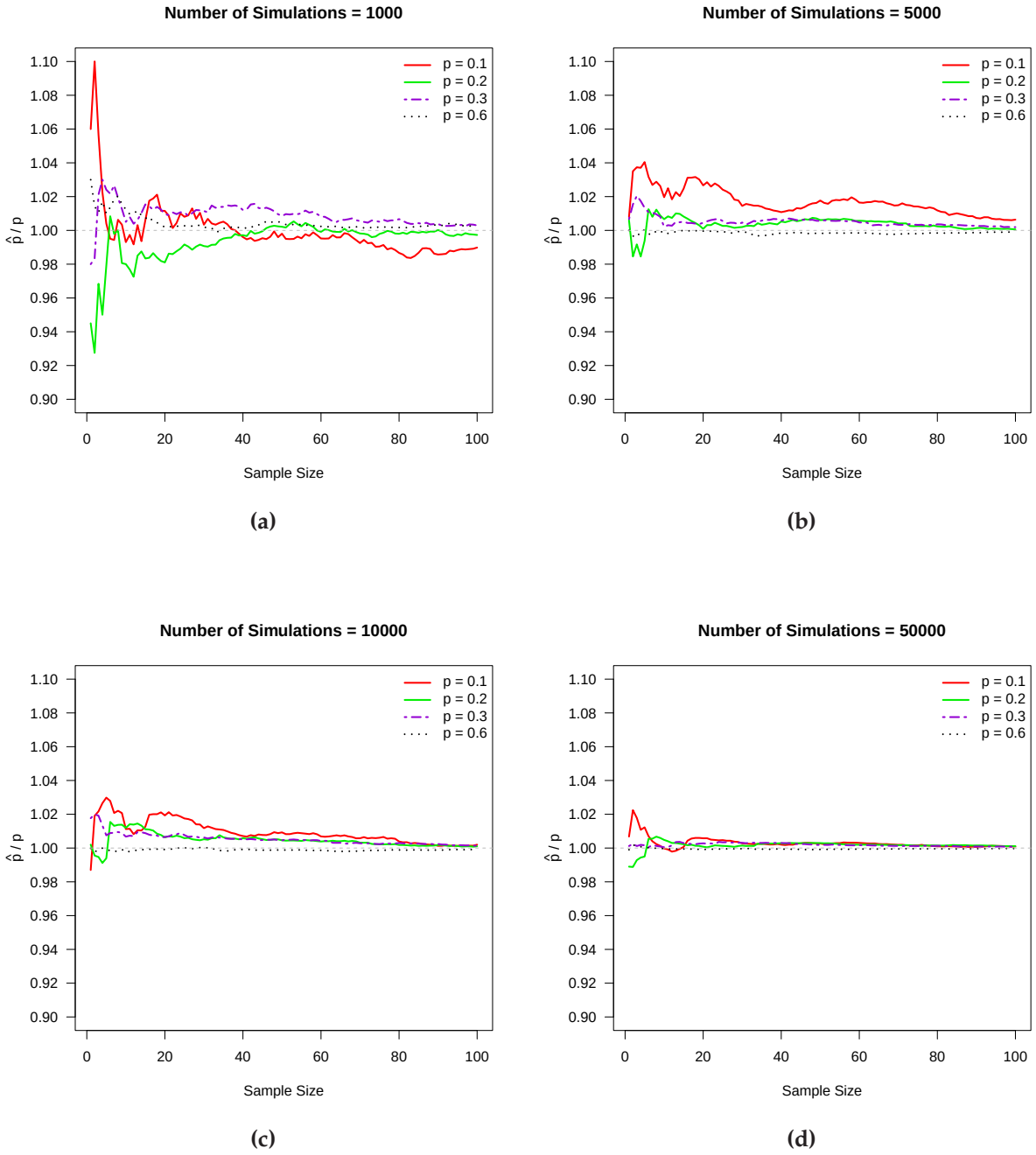
Since characteristics of ZIP data may change with the magnitude of λ, we also explored the case of $\lambda = 10$, keeping all other conditions the same as for $\lambda = 2$. Table 12.2 presents the simulated results for $\lambda = 10$. Figure 12.4 is the visual representation of Table 12.2. The significant difference between cases $\lambda = 2$ and $\lambda = 10$ is that for $\lambda = 10$ there is no apparent variation of $\hat{p}$ when sample sizes are small, but the accuracy of $\hat{p}$ is still an issue, especially when the number of simulation is not sufficiently large, as demonstrated in Figure 12.4(a).

Information provided in these tables and figures indicates that $\hat{p}$ calculated from small sample sizes (say less than 6) is unreliable, even when the number of simulations is sufficiently large. As one of key components of the GLR chart, the MLE with poor performance will result in an increased control limit for a given in-control chart performance, leading to poor surveillance performance of the GLR chart. Furthermore, it is well known that simulations for the GLR algorithm are time-consuming relative to other methods. As a consequence, it is necessary to make some modification of $\hat{p}$ when sample sizes are small.

To improve the accuracy of $\hat{p}$, we explored the Bayesian data analysis using Markov chain Monte Carlo (MCMC) methods, in which the prior distribution of p is assumed to be a beta distribution. Preliminary results indicate the value of $\hat{p}$, i.e., the mean of the posterior distribution, mainly depends on the parameter settings of the beta distribution.

Table 12.2: Effect of small sample sizes on the MLE for p in p-GLR charts, where $\lambda = 10$. The values in this table are of $\hat{p}/p$

	N.Simulation $= 1,000$				N.Simulation $= 5,000$				N.Simulation $= 10,000$				N.Simulation $= 50,000$			
Sample Size	$p = 0.1$	0.2	0.3	0.6	0.1	0.2	0.3	0.6	0.1	0.2	0.3	0.6	0.1	0.2	0.3	0.6
1	1.060	0.945	0.980	1.030	1.008	1.006	1.007	1.000	0.987	1.002	1.018	1.000	1.007	0.989	1.001	0.999
2	1.100	0.928	0.983	1.014	1.035	0.985	1.015	0.997	1.019	0.996	1.020	0.998	1.022	0.989	1.002	1.002
3	1.057	0.968	1.021	1.011	1.037	0.992	1.020	0.996	1.022	0.995	1.019	0.999	1.018	0.993	1.001	1.002
4	1.023	0.950	1.030	1.017	1.037	0.985	1.017	1.000	1.027	0.991	1.013	1.000	1.011	0.994	1.002	1.001
5	1.004	0.979	1.024	1.010	1.040	0.994	1.013	1.000	1.030	0.994	1.008	0.999	1.012	0.995	1.002	1.001
6	0.995	1.008	1.022	1.014	1.032	1.013	1.011	0.998	1.028	1.015	1.009	0.998	1.007	1.005	1.000	1.000
7	0.994	0.998	1.027	1.018	1.027	1.008	1.011	0.997	1.021	1.013	1.009	0.998	1.004	1.006	1.002	1.000
8	1.006	1.000	1.019	1.020	1.029	1.012	1.010	1.000	1.022	1.014	1.010	0.999	1.002	1.007	1.001	1.000
9	1.003	0.981	1.013	1.016	1.026	1.009	1.008	1.000	1.021	1.014	1.009	0.999	1.001	1.006	1.001	1.000
10	0.993	0.980	1.005	1.012	1.020	1.007	1.002	0.998	1.011	1.012	1.007	0.997	1.000	1.005	1.000	1.000
11	0.997	0.977	1.008	1.010	1.025	1.008	1.003	0.998	1.011	1.014	1.007	0.998	0.999	1.004	1.001	1.000
12	0.992	0.973	1.004	1.010	1.018	1.007	1.003	0.999	1.008	1.014	1.007	0.998	0.998	1.003	1.002	0.999
13	1.003	0.985	1.006	1.012	1.023	1.010	1.005	0.999	1.011	1.015	1.009	0.999	0.998	1.003	1.004	1.000
14	0.994	0.988	1.010	1.009	1.020	1.010	1.005	1.000	1.010	1.013	1.009	0.999	1.000	1.003	1.004	0.999
15	1.006	0.983	1.015	1.007	1.025	1.008	1.005	1.000	1.013	1.011	1.009	0.999	1.001	1.002	1.003	0.999
16	1.018	0.984	1.014	1.008	1.031	1.007	1.004	1.000	1.020	1.011	1.008	0.999	1.004	1.002	1.003	0.999
17	1.019	0.987	1.012	1.005	1.031	1.006	1.004	1.000	1.020	1.010	1.007	0.999	1.006	1.002	1.002	0.999
18	1.021	0.984	1.014	1.005	1.032	1.004	1.004	1.000	1.020	1.008	1.007	0.999	1.006	1.001	1.002	1.000
19	1.012	0.982	1.012	1.003	1.030	1.003	1.004	0.999	1.021	1.008	1.006	0.999	1.006	1.001	1.003	0.999
20	1.012	0.981	1.010	1.002	1.027	1.001	1.005	1.000	1.019	1.006	1.007	0.999	1.006	1.001	1.003	0.999
21	1.009	0.986	1.009	1.002	1.029	1.003	1.005	0.999	1.021	1.007	1.007	0.999	1.006	1.000	1.003	0.999
22	1.001	0.986	1.011	1.003	1.026	1.003	1.006	1.000	1.019	1.007	1.008	0.999	1.005	1.001	1.003	0.999
23	1.004	0.988	1.010	1.003	1.028	1.005	1.007	1.000	1.020	1.007	1.008	1.000	1.005	1.002	1.002	1.000
24	1.010	0.989	1.011	1.002	1.026	1.004	1.007	0.999	1.019	1.007	1.009	1.000	1.005	1.001	1.004	0.999
25	1.008	0.992	1.011	1.003	1.024	1.003	1.006	0.999	1.018	1.006	1.008	1.000	1.005	1.001	1.003	0.999
26	1.009	0.990	1.009	1.003	1.023	1.003	1.004	0.999	1.017	1.006	1.007	1.000	1.004	1.001	1.003	1.000
27	1.013	0.989	1.010	1.003	1.022	1.002	1.005	0.998	1.016	1.005	1.007	1.000	1.004	1.001	1.003	0.999
28	1.007	0.990	1.012	1.003	1.018	1.002	1.004	0.999	1.014	1.005	1.007	1.000	1.004	1.001	1.003	1.000
29	1.010	0.992	1.012	1.003	1.018	1.002	1.004	0.999	1.014	1.005	1.006	1.000	1.004	1.001	1.003	1.000
30	1.003	0.991	1.011	1.002	1.015	1.002	1.005	0.999	1.012	1.005	1.006	1.000	1.003	1.001	1.003	1.000



(a) **(b)**

(c) **(d)**

Figure 12.4: Effect of small sample sizes on the MLE for p in p-GLR charts, where $\lambda = 10$. (a) number of simulations $= 1,000$; (b) number of simulations $= 5,000$; (c) number of simulations $= 10,000$; (d) number of simulations $= 50,000$.

More sophisticated priors may improve the performance of the Bayesian method, but this will lead to the GLR chart being even more complicated and less practical. A simple solution to the MLE issue due to small sizes resulting from the change point framework will be discussed in Section 12.3.3.

In the following section, the GLR chart statistics will be derived based on the MLE and the changepoint framework.

12.3.2 Statistics of p-GLR charts

Suppose a sequential sample $\{X_1, X_2, \cdots, X_k\}$ is obtained and the value of p has changed from p_0 to an unknown value p_1 since time point τ^*, where $\tau < \tau^* < \tau + 1 \leq k$. If we let $\beta_{k,\tau}(p_1)$ be the log likelihood ratio from the change time point τ^* to the end time point k, then $\beta_{k,\tau}(p_1)$ can be expressed as

$$\begin{aligned}
\beta_{k,\tau}(p_1) &= \sum_{i=\tau+1}^{k} \ln \frac{f(x_i; p_1^{k,\tau}, \lambda)}{f(x_i; p_0, \lambda)} \\
&= \sum_{j=0}^{l} n_j^{k,\tau} \ln \frac{f(j; p_1^{k,\tau}, \lambda)}{f(j; p_0, \lambda)} \\
&= n_0^{k,\tau} \ln \frac{1 + (e^{-\lambda} - 1)p_1^{k,\tau}}{1 + (e^{-\lambda} - 1)p_0} + (n^{k,\tau} - n_0^{k,\tau}) \ln \frac{p_1^{k,\tau}}{p_0},
\end{aligned} \tag{12.11}$$

where $n^{k,\tau}$, $n_0^{k,\tau}$, j, and $n_j^{k,\tau}$ share the same definition as n, n_0, j, and n_j respectively in Equation (12.7). The superscript (k, τ) is used to indicate all these parameters are related to changepoint τ^* and endpoint k. In this sense, the MLE for p subject to the condition (k, τ) is written as

$$\hat{p}_1^{k,\tau} = \min\left(1, \quad \frac{n^{k,\tau} - n_0^{k,\tau}}{n^{k,\tau}} \cdot \frac{1}{1 - e^{-\lambda}}\right). \tag{12.12}$$

Hence, the p-GLR chart statistic which is doubly maximized in terms of τ and p_1 at time point k can be given by

$$\begin{aligned} R_k^p &= \max_{0\le\tau<k} \sup_{p_1} \beta_{k,\tau}(p_1) \\ &= \max_{0\le\tau<k} \Big\{ n_0^{k,\tau} \ln \frac{1+(e^{-\lambda}-1)\hat{p}_1^{k,\tau}}{1+(e^{-\lambda}-1)p_0} + (n^{k,\tau}-n_0^{k,\tau}) \ln \frac{\hat{p}_1^{k,\tau}}{p_0} \Big\}. \end{aligned} \tag{12.13}$$

As discussed in Section 7.1, if the value of k is very large, the computational work to obtain R_k^p will be time-consuming. An approach to this concern is to restrict the calculation within a window with size m, and then the statistic becomes

$$R_{m,k}^p = \begin{cases} \max\limits_{0\le\tau<k} \Big\{ n_0^{k,\tau} \ln \frac{1+(e^{-\lambda}-1)\hat{p}_1^{k,\tau}}{1+(e^{-\lambda}-1)p_0} + (n^{k,\tau}-n_0^{k,\tau}) \ln \frac{\hat{p}_1^{k,\tau}}{p_0} \Big\} & \text{if } k = 1, 2, \cdots, m \\ \max\limits_{k-m\le\tau<k} \Big\{ n_0^{k,\tau} \ln \frac{1+(e^{-\lambda}-1)\hat{p}_1^{k,\tau}}{1+(e^{-\lambda}-1)p_0} + (n^{k,\tau}-n_0^{k,\tau}) \ln \frac{\hat{p}_1^{k,\tau}}{p_0} \Big\} & \text{if } k = m+1, m+2, \cdots. \end{cases} \tag{12.14}$$

Equation (12.13) is derived under the assumption that there is a shift in p, which means the shift can be either upward or downward, so the resulting GLR statistic is a two-sided statistic. In practice, one may be mainly concerned with increases in the number of nonzero counts, so one-sided statistics are desired. Using a sign function, we can modify two-sided statistics to one-sided statistics. If we let $R_k^{p,s}$ be the one-sided statistic corresponding to the two-sided statistic R_k^p, then the modification can be expressed as

$$\begin{aligned} R_k^{p,s} &= \operatorname{sgn}(\hat{p}_1^{k,\tau} - p_0) \cdot R_k^p \\ &= \operatorname{sgn}(\hat{p}_1^{k,\tau} - p_0) \cdot \max_{0\le\tau<k} \Big\{ n_0^{k,\tau} \ln \frac{1+(e^{-\lambda}-1)\hat{p}_1^{k,\tau}}{1+(e^{-\lambda}-1)p_0} + (n^{k,\tau}-n_0^{k,\tau}) \ln \frac{\hat{p}_1^{k,\tau}}{p_0} \Big\}. \end{aligned} \tag{12.15}$$

If $R_k^{p,s}$ is maximized at time point $\hat{\tau}$, then Equation (12.15) will be reduced to

$$R_k^{p,s} = \text{sgn}(\hat{p}_1^{k,\hat{\tau}} - p_0) \cdot \left[n_0^{k,\hat{\tau}} \ln \frac{1 + (e^{-\lambda} - 1)\hat{p}_1^{k,\hat{\tau}}}{1 + (e^{-\lambda} - 1)p_0} + (n^{k,\hat{\tau}} - n_0^{k,\hat{\tau}}) \ln \frac{\hat{p}_1^{k,\hat{\tau}}}{p_0} \right]. \tag{12.16}$$

Figure 12.6 visually illustrates how the sign function changes the two-sided p-GLR chart statistics to one-sided statistics for a specific case where $p_0 = 0.2$ and $\lambda = 2$; see Section 7.3 and Appendix A for more discussion on the working mechanism and the benefits of using the sign function.

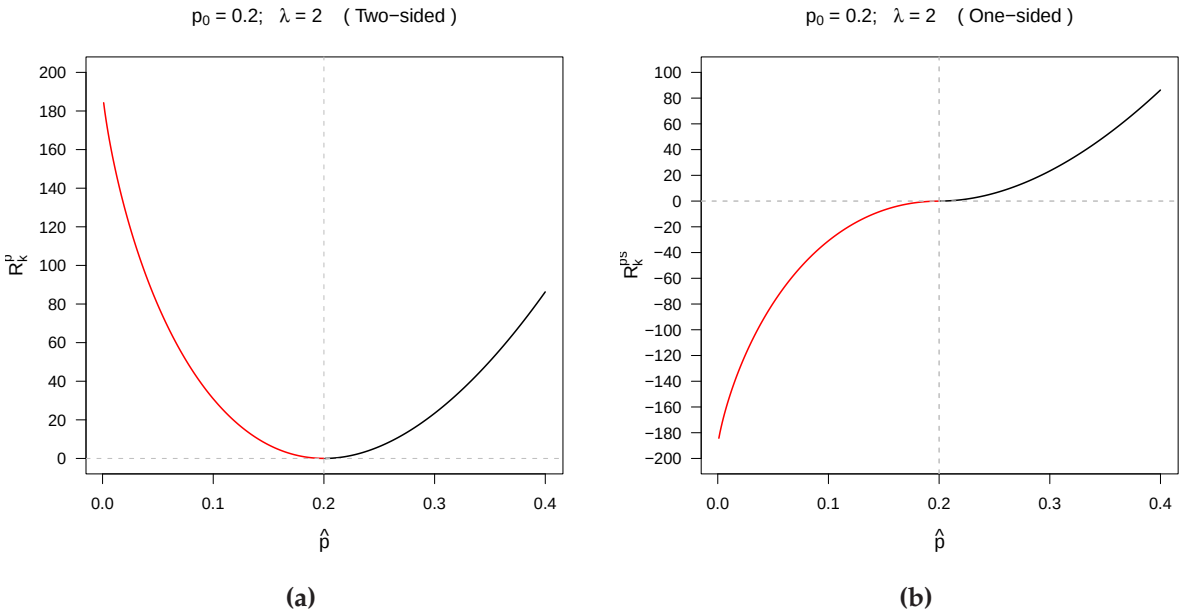


Figure 12.5: Relationship between the GLR statistics and the maximum likelihood estimators for p in a p-GLR chart, where $p_0 = 0.2$, $\lambda = 2$. (a) two-sided statistics; (b) one-sided statistics.

12.3.3 A modified scheme for the change point framework

As discussed in Section 12.3.1, the GLR algorithm assumes each past time point can be a change point, which may result in a bias of the MLE for p and consequently produce

poor surveillance performance. A simple solution to this issue can be to ignore some small sample sizes which are used for calculating the MLE.

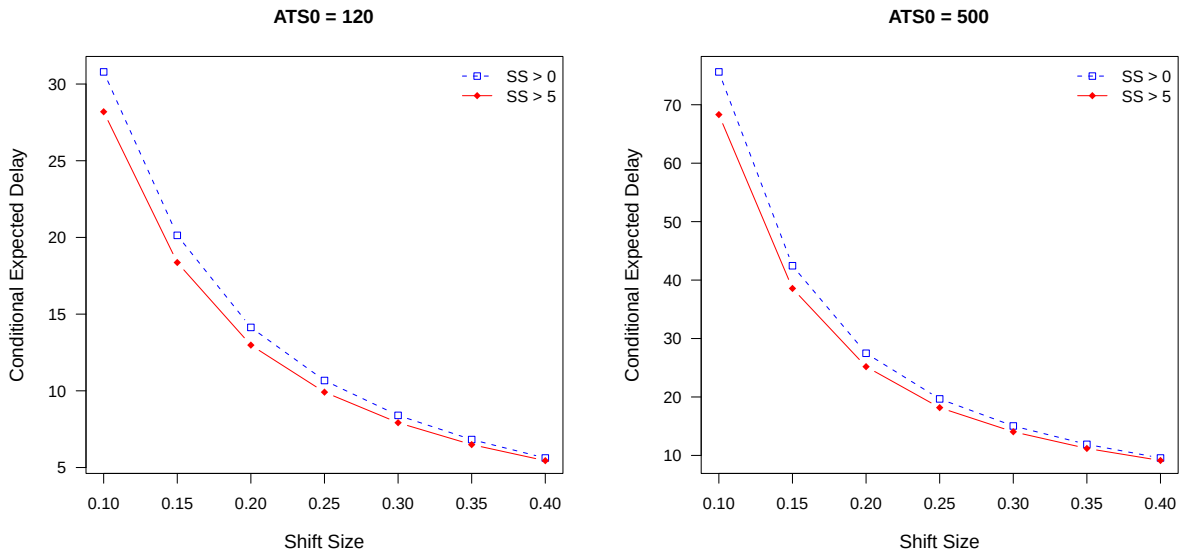
Table 12.3 provides a comparison of conditional expected delay (CED) values for all sample sizes used and only sample sizes greater than five considered, where $p_0 = 0.2$, $p = 6$, and SS stands for the sample size. Shift sizes in the first column are the absolute values of shifts in p, and other values in this table are CED values. For example, if the shift size is 0.1, then the value of p in the out-of-control condition is 0.3. Figure 12.6 intuitively displays the results in Table 12.3. It can be seen that the conditional expected delay values of only considering sample sizes greater than five are lower than those of considering all sample sizes for each shift size, which means the modified change point framework has improved the surveillance performance.

Table 12.3: A comparison of conditional expected delay values of p-GLR charts when all sample sizes used and ignoring some small sample sizes, where $p_0 = 0.2$ and $\lambda = 6$

	$\text{ATS}_0 = 120$		$\text{ATS}_0 = 500$	
Shift size (δ)	$SS > 0$	$SS > 5$	$SS > 0$	$SS > 5$
0.10	30.79	28.19	75.61	68.31
0.15	20.13	18.36	42.45	38.57
0.20	14.13	12.98	27.48	25.18
0.25	10.67	9.913	19.66	18.17
0.30	8.39	7.921	15.03	14.04
0.35	6.82	6.493	11.89	11.19
0.40	5.62	5.446	9.55	9.13

12.4 Derivation of λ-GLR chart statistics

The preceding section proposed the statistic for detecting shifts in the parameter p. In this section, we will develop a GLR chart called the λ-GLR chart, which is designed to



(a) **(b)**

Figure 12.6: Comparison of CED values of considering all sample sizes and sample sizes greater than 5 for p-GLR charts, where $p_0 = 0.2$, $\lambda = 6$ (a) $\text{ATS}_0 = 120$; (b) $\text{ATS}_0 = 500$.

monitor the parameter λ.

12.4.1 MLE for λ-GLR charts

Suppose the parameter p remains unchanged over a ZIP process, then based on Equation (12.7) the score equation for the λ-GLR chart is represented by

$$\frac{\partial \ln L}{\partial \lambda} = -\frac{n_0 p e^{-\lambda}}{1 - p + p e^{-\lambda}} - (n - n_0)\left(1 - \frac{\overline{x}^+}{\lambda}\right), \tag{12.17}$$

and the resulting MLE for λ can be expressed as

$$\frac{p}{p + (1-p)e^{\hat{\lambda}}} = \frac{n - n_0}{n_0}\left(\frac{\overline{x}^+}{\hat{\lambda}} - 1\right). \tag{12.18}$$

Since $0 \leq p \leq 1$, it is easy to derive the restriction $\frac{n-n_0}{n}\overline{x}^+ \leq \hat{\lambda} \leq \overline{x}^+$. However, solving for $\hat{\lambda}$ is not straightforward, and suitable numerical methods are required. Derivation and discussion of the MLE for λ-GLR charts are given in Appendix D.

If the MLE for λ is derived from a sample with small size, the estimation may be unstable as illustrated for $\hat{p}$ in the preceding sections. In this section, the effect of sample size on the MLE $\hat{\lambda}$ will be explored.

We performed studies to investigate the MLE performance under various settings, such as varying levels of p, different in-control values of λ, and different simulation numbers. Table 12.4 summarizes the numerical results based on 1000 simulations, and Figure 12.7 displays the information. This figure indicates that the MLE performance becomes worse as the magnitude of λ decreases. When the values of p and λ are small, poor performance is apparent as indicated in Figure 12.4(a). Specifically, there still exists considerable variation even when sample sizes are as large as 100. Taking a closer examination of the $\lambda = 2$ case, we can see overestimation for small sample sizes and underestimation for relatively large sample sizes. Compared with the performance of the MLE for p, the effect of small sample sizes on the performance of MLE for λ is not significant when the λ value is not too small. However, the overall performance of $\hat{\lambda}$ appears not as good as that of $\hat{p}$. Table 12.5 and Figure 12.8 present the corresponding results when the number of simulations is 10,000. Figures 12.7 and 12.8 indicate that the MLE performance based on 1000 simulations is similar to that based on more simulations.

12.4.2 Statistics of λ-GLR charts

If the MLE for λ is available, then the λ-GLR chart statistic can be developed based on log likelihood ratios. We assume the value of λ has been shifted from λ_0 to an unknown

Table 12.4: Effect of small sample sizes on the MLE for λ in λ-GLR charts, where the number of simulations is 1,000. The values in this table are of $\hat{\lambda}/\lambda$

	$p_0 = 0.1$				$p_0 = 0.2$				$p_0 = 0.3$				$p_0 = 0.6$			
Sample Size	$\lambda_0 = 2$	4	6	10	2	4	6	10	2	4	6	10	2	4	6	10
1	1.009	1.005	0.997	1.003	1.027	1.001	1.003	0.997	1.042	0.985	0.993	0.985	1.096	0.994	0.988	0.988
2	1.020	1.010	0.994	1.010	1.047	0.997	0.999	0.997	1.065	0.982	0.993	0.986	1.059	0.970	0.995	0.999
3	1.029	1.010	0.992	1.012	1.051	0.985	1.001	0.996	1.052	0.987	0.991	0.980	1.031	0.964	1.001	1.002
4	1.038	1.007	0.995	1.012	1.031	0.986	1.001	0.997	1.044	0.978	0.997	0.983	0.996	0.970	1.001	1.002
5	1.041	1.009	0.992	1.008	1.023	0.986	1.006	0.996	1.037	0.973	0.990	0.982	0.998	0.971	0.996	1.004
6	1.036	1.008	0.997	1.004	1.020	0.995	1.007	0.994	1.021	0.976	0.984	0.987	0.998	0.980	0.989	0.998
7	1.043	1.002	0.998	0.999	1.013	0.990	1.006	1.002	1.016	0.973	0.978	0.990	0.990	0.984	0.989	1.000
8	1.037	1.001	0.999	0.996	1.001	0.984	1.003	1.001	1.000	0.969	0.977	0.993	0.997	0.980	0.987	1.001
9	1.035	1.008	0.998	0.998	0.999	0.975	1.004	1.002	0.997	0.967	0.977	0.996	0.992	0.977	0.989	1.001
10	1.033	1.003	1.000	0.997	0.995	0.974	1.010	0.999	0.984	0.968	0.979	0.999	0.995	0.980	0.993	1.002
11	1.036	0.994	1.001	0.998	0.990	0.974	1.008	1.001	0.977	0.967	0.982	1.000	0.991	0.982	0.992	1.001
12	1.028	0.994	0.998	0.998	0.997	0.973	1.009	0.999	0.977	0.971	0.981	1.002	0.991	0.986	0.993	1.000
13	1.025	0.994	0.995	1.001	0.990	0.980	1.007	0.994	0.971	0.971	0.984	1.006	0.988	0.989	0.994	1.001
14	1.022	0.991	0.991	1.001	0.982	0.974	1.006	0.994	0.969	0.970	0.982	1.004	0.990	0.993	0.995	1.000
15	1.014	0.991	0.990	1.004	0.986	0.975	1.004	0.994	0.964	0.975	0.988	1.004	0.991	0.992	0.996	1.000
16	1.008	0.992	0.992	1.001	0.981	0.978	0.999	0.996	0.966	0.980	0.988	1.001	0.990	0.991	0.996	0.999
17	1.002	0.993	0.997	1.000	0.982	0.973	1.001	0.997	0.966	0.984	0.992	1.000	0.989	0.992	0.997	1.000
18	0.994	0.991	0.997	0.998	0.975	0.975	1.002	0.998	0.966	0.986	0.989	1.001	0.988	0.992	0.998	1.001
19	0.993	0.989	0.996	0.996	0.984	0.973	1.002	0.998	0.967	0.990	0.990	1.001	0.987	0.994	1.000	1.000
20	0.988	0.994	0.997	0.999	0.976	0.971	1.000	0.999	0.969	0.991	0.992	1.000	0.986	0.994	1.001	1.000
21	0.989	0.986	0.997	0.997	0.975	0.979	0.996	1.001	0.966	0.990	0.993	1.001	0.986	0.992	1.001	1.001
22	0.997	0.984	0.998	0.999	0.974	0.978	0.996	1.002	0.965	0.990	0.995	1.002	0.984	0.993	1.002	1.000
23	0.997	0.982	0.997	0.997	0.969	0.976	0.995	1.004	0.967	0.991	0.995	1.002	0.986	0.994	1.002	1.000
24	0.992	0.980	1.001	0.996	0.969	0.978	0.994	1.003	0.964	0.994	0.996	1.002	0.987	0.994	1.002	1.001
25	0.996	0.982	0.999	0.994	0.971	0.980	0.995	1.002	0.967	0.993	0.995	1.002	0.986	0.995	1.002	1.000
26	0.990	0.980	0.998	0.995	0.973	0.982	0.996	1.001	0.969	0.993	0.993	1.001	0.987	0.994	1.001	1.001
27	0.987	0.981	0.996	0.995	0.974	0.986	0.998	1.000	0.969	0.992	0.992	1.001	0.988	0.994	1.000	1.002
28	0.983	0.982	0.998	0.993	0.974	0.988	0.998	1.000	0.968	0.991	0.993	1.000	0.986	0.994	0.999	1.001
29	0.983	0.980	0.997	0.993	0.973	0.988	0.998	1.001	0.968	0.992	0.994	1.000	0.985	0.993	0.998	1.000
30	0.979	0.979	0.996	0.993	0.974	0.990	0.999	1.000	0.969	0.990	0.993	1.001	0.986	0.994	0.998	1.000

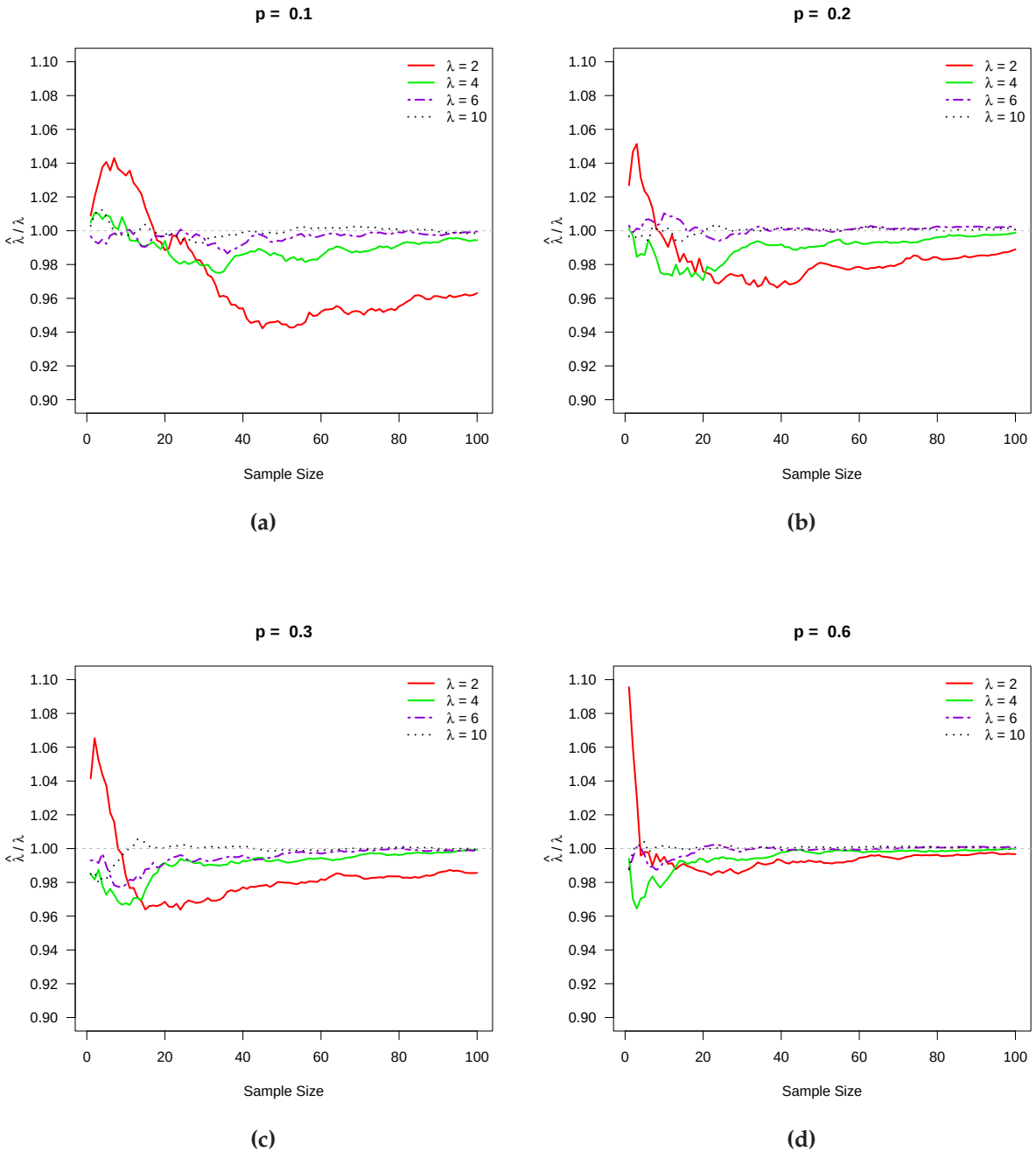
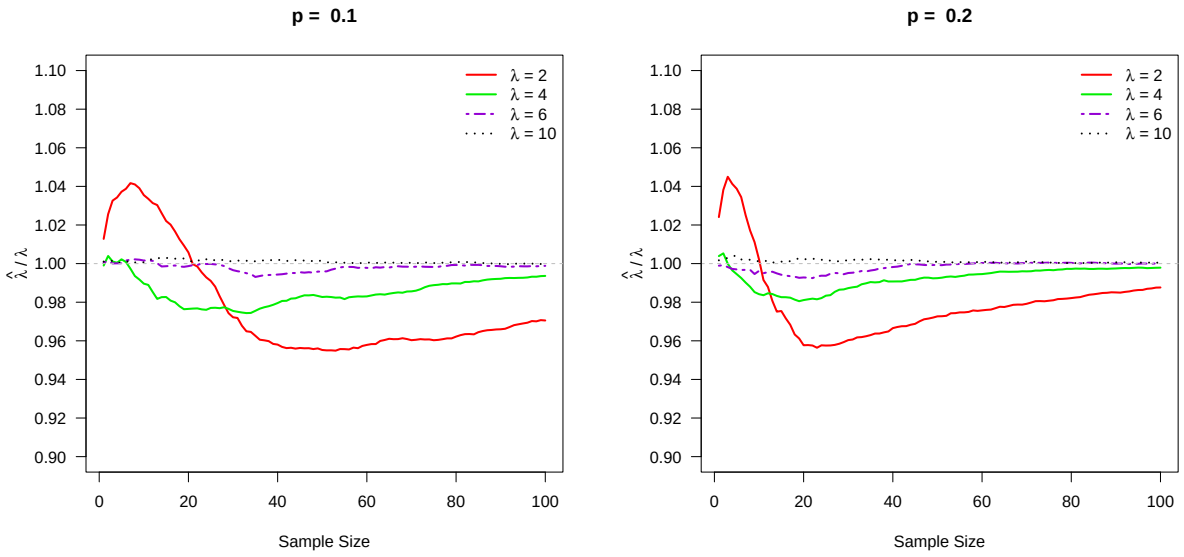


Figure 12.7: Effect of small sample sizes on the MLE for λ in λ-GLR charts, where the number of simulations is 1,000. (a) $p = 0.1$; (b) $p = 0.2$; (c) $p = 0.3$; (d) $p = 0.6$.

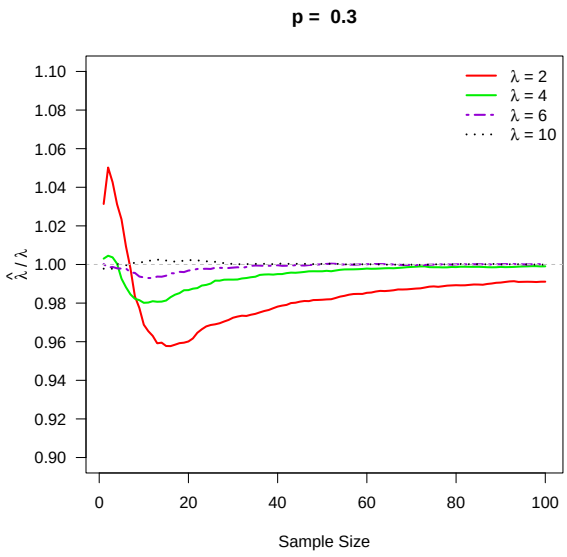
Table 12.5: Effect of small sample sizes on the MLE for λ in λ-GLR charts, where the number of simulations is 10,000. The values in this table are of $\hat{\lambda}/\lambda$

	$p_0 = 0.1$				$p_0 = 0.2$				$p_0 = 0.3$				$p_0 = 0.6$			
Sample Size	$\lambda_0 = 2$	4	6	10	2	4	6	10	2	4	6	10	2	4	6	10
1	1.013	0.999	1.001	1.001	1.024	1.004	0.999	1.002	1.031	1.003	1.000	0.998	1.082	1.004	1.004	0.997
2	1.026	1.004	1.001	1.001	1.038	1.005	0.999	1.002	1.050	1.005	0.998	0.998	1.057	0.991	1.000	0.997
3	1.033	1.001	1.000	1.001	1.045	1.000	0.998	1.003	1.043	1.004	0.999	0.998	1.016	0.980	0.997	0.997
4	1.034	1.000	1.000	1.000	1.041	0.996	0.997	1.004	1.031	1.000	0.998	1.000	0.988	0.981	0.996	0.999
5	1.037	1.002	1.000	1.001	1.039	0.995	0.997	1.004	1.023	0.993	0.998	1.000	0.976	0.980	0.992	1.000
6	1.039	1.001	1.002	1.001	1.034	0.993	0.997	1.002	1.009	0.988	0.998	1.000	0.974	0.983	0.993	1.000
7	1.042	0.997	1.002	1.002	1.025	0.990	0.997	1.002	0.998	0.985	0.996	1.001	0.974	0.985	0.996	1.001
8	1.041	0.994	1.002	1.001	1.017	0.988	0.997	1.002	0.983	0.982	0.996	1.001	0.976	0.986	0.998	1.000
9	1.039	0.992	1.002	1.001	1.011	0.985	0.995	1.002	0.977	0.982	0.994	1.001	0.978	0.988	0.998	1.000
10	1.035	0.990	1.002	1.001	1.002	0.984	0.996	1.001	0.969	0.980	0.993	1.001	0.979	0.991	0.999	1.000
11	1.034	0.989	1.002	1.001	0.992	0.984	0.995	1.001	0.965	0.980	0.993	1.002	0.980	0.992	0.999	1.000
12	1.031	0.985	1.001	1.003	0.988	0.985	0.996	1.002	0.963	0.981	0.993	1.002	0.980	0.994	0.999	1.000
13	1.030	0.982	1.000	1.003	0.981	0.984	0.996	1.000	0.959	0.981	0.994	1.003	0.982	0.994	1.000	1.000
14	1.026	0.983	0.999	1.003	0.975	0.983	0.995	1.001	0.960	0.981	0.994	1.002	0.984	0.994	1.001	1.001
15	1.022	0.983	0.999	1.003	0.975	0.983	0.994	1.001	0.958	0.981	0.994	1.002	0.984	0.994	1.000	1.000
16	1.020	0.981	0.999	1.003	0.972	0.983	0.994	1.001	0.958	0.983	0.995	1.002	0.985	0.995	1.000	1.000
17	1.017	0.980	0.999	1.003	0.968	0.982	0.993	1.001	0.958	0.984	0.995	1.001	0.984	0.995	1.000	1.000
18	1.013	0.978	0.999	1.003	0.963	0.981	0.993	1.002	0.959	0.985	0.996	1.002	0.986	0.996	1.001	1.000
19	1.009	0.976	0.998	1.003	0.961	0.981	0.993	1.003	0.959	0.987	0.996	1.002	0.986	0.996	1.001	1.000
20	1.006	0.977	0.999	1.002	0.958	0.981	0.993	1.002	0.960	0.987	0.997	1.002	0.987	0.996	1.001	1.000
21	1.000	0.977	0.999	1.001	0.958	0.981	0.993	1.002	0.962	0.987	0.997	1.002	0.987	0.996	1.001	1.000
22	0.998	0.977	0.999	1.002	0.958	0.982	0.992	1.002	0.965	0.988	0.998	1.002	0.988	0.996	1.001	1.000
23	0.995	0.976	1.000	1.001	0.956	0.982	0.993	1.002	0.966	0.989	0.998	1.002	0.989	0.997	1.001	1.000
24	0.993	0.976	1.000	1.002	0.958	0.982	0.994	1.002	0.968	0.989	0.998	1.002	0.990	0.997	1.001	1.000
25	0.991	0.977	1.000	1.001	0.958	0.983	0.994	1.001	0.969	0.990	0.998	1.002	0.990	0.997	1.001	1.000
26	0.988	0.977	1.000	1.002	0.958	0.984	0.995	1.002	0.969	0.992	0.998	1.001	0.990	0.997	1.001	1.000
27	0.983	0.977	0.999	1.002	0.958	0.985	0.995	1.001	0.970	0.992	0.998	1.001	0.991	0.997	1.001	1.000
28	0.979	0.977	0.999	1.002	0.958	0.986	0.995	1.001	0.970	0.992	0.998	1.001	0.991	0.998	1.000	1.000
29	0.974	0.977	0.998	1.001	0.959	0.987	0.995	1.002	0.971	0.992	0.998	1.000	0.992	0.998	1.000	1.000
30	0.972	0.975	0.997	1.001	0.960	0.987	0.995	1.002	0.972	0.992	0.998	1.000	0.992	0.998	1.000	1.000

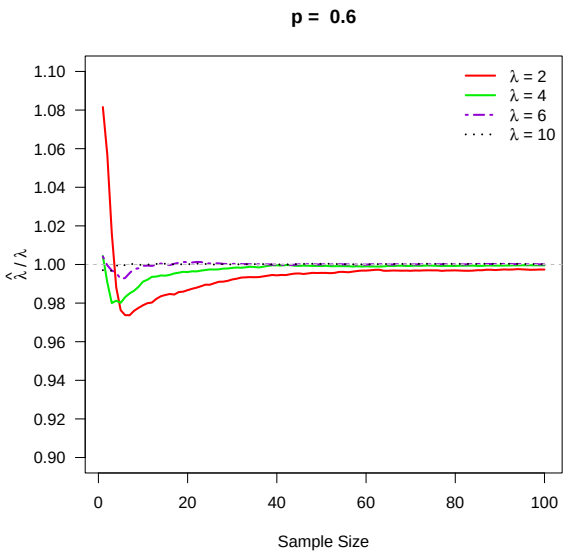


(a)

(b)



(c)



(d)

Figure 12.8: Effect of small sample sizes on the MLE for λ in λ-GLR charts, where the number of simulations is 10,000. (a) $p = 0.1$; (b) $p = 0.2$; (c) $p = 0.3$; (d) $p = 0.6$.

value of λ_1 since time point τ^*, while the value of p remains unchanged.

If we let $\beta_{k,\tau}(\lambda_1)$ be the log likelihood ratio of the λ-GLR chart from the changepoint τ^* to the end time point k, with similar definitions and notation to those for the p-GLR chart, then $\beta_{k,\tau}(\lambda_1)$ can be expressed by

$$\begin{aligned}\beta_{k,\tau}(\lambda_1) &= \sum_{j=0}^{l} n_j^{k,\tau} \ln \frac{f(j;\lambda_1^{k,\tau},p)}{f(j;\lambda_0,p)} \\ &= n_0^{k,\tau} \ln \frac{1+(e^{-\lambda_1^{k,\tau}}-1)p}{1+(e^{-\lambda_0}-1)p} + (n^{k,\tau}-n_0^{k,\tau})\Big[\overline{x}_{k,\tau}^{+} \ln \frac{\lambda_1^{k,\tau}}{\lambda_0} - (\lambda_1^{k,\tau}-\lambda_0)\Big]. \end{aligned} \quad (12.19)$$

The statistics of the λ-GLR charts are thus calculated as

$$\begin{aligned} R_k^{\lambda} &= \max_{0\le\tau<k} \sup_{\lambda_1} \beta_{k,\tau}(\lambda_1) \\ &= \max_{0\le\tau<k} \Big\{ n_0^{k,\tau} \ln \frac{1+(e^{-\hat{\lambda}_1^{k,\tau}}-1)p}{1+(e^{-\lambda_0}-1)p} + (n^{k,\tau}-n_0^{k,\tau})\Big[\overline{x}_{k,\tau}^{+} \ln \frac{\hat{\lambda}_1^{k,\tau}}{\lambda_0} - (\hat{\lambda}_1^{k,\tau}-\lambda_0)\Big]\Big\}, \end{aligned} \quad (12.20)$$

and the window-limited statistics are

$$R_{m,k}^{\lambda} = \begin{cases} \max\limits_{0\le\tau<k} \Big\{ n_0^{k,\tau} \ln \dfrac{1+(e^{-\hat{\lambda}_1^{k,\tau}}-1)p}{1+(e^{-\lambda_0}-1)p} \\ \qquad +(n^{k,\tau}-n_0^{k,\tau})\Big[\overline{x}_{k,\tau}^{+} \ln \dfrac{\hat{\lambda}_1^{k,\tau}}{\lambda_0} - (\hat{\lambda}_1^{k,\tau}-\lambda_0)\Big]\Big\} & \text{if } k = 1,2,\cdots,m \\ \max\limits_{k-m\le\tau<k} \Big\{ n_0^{k,\tau} \ln \dfrac{1+(e^{-\hat{\lambda}_1^{k,\tau}}-1)p}{1+(e^{-\lambda_0}-1)p} \\ \qquad +(n^{k,\tau}-n_0^{k,\tau})\Big[\overline{x}_{k,\tau}^{+} \ln \dfrac{\hat{\lambda}_1^{k,\tau}}{\lambda_0} - (\hat{\lambda}_1^{k,\tau}-\lambda_0)\Big]\Big\} & \text{if } k = m+1,m+2,\cdots. \end{cases} \quad (12.21)$$

It is easy to show that when $p = 1$ Equation (12.20) reduces to Equation (7.5), which is the expression for the GLR chart statistics for Poisson data. This is not surprising, because the Poisson distributed data is a special case of the ZIP distributed data.

Since R_k^{λ} in Equation (12.20) is derived under the assumption that there is a shift in

λ, the values of R_k^λ are the two-sided statistics for λ-GLR charts. As illustrated before, a sign function can modify this two-sided statistic to one-sided form. This modification can be expressed as

$$R_k^{\lambda,s} = \text{sgn}(\hat{\lambda}_1^{k,\tau} - \lambda_0) \cdot R_k^\lambda. \tag{12.22}$$

12.5 Development of p-λ-GLR charts

In the preceding sections, we derived control chart statistics for individual GLR charts, which are designed to detect shifts either in p or in λ. However, in the biosurveillance context, it is more often that both parameters can shift during disease outbreaks. In this sense, a control chart for detecting shifts in these two parameters simultaneously is desired. In the framework of statistical process control, a control chart monitoring two process parameters simultaneously can be either consist of a single chart or a combination of two control charts. We briefly introduce a combination method to simultaneously monitor p and λ in this section. A single chart for this purpose will be given in the next section.

Suppose the p-GLR chart and the λ-GLR chart have been developed, then we can combine these two GLR charts into a new GLR chart called the p-λ-GLR chart, in which the p-GLR chart and the λ-GLR chart are designed to detect shifts in p and λ, respectively. In this sense, the p-λ-GLR chart statistics are therefore the combination of $R_k^{p,s}$ and $R_k^{\lambda,s}$, as proposed in preceding sections. The p-λ-GLR chart signals if one of these two charts signals. The control limits of the p-λ-GLR chart appear in pairs, in which one is the threshold for the p-GLR chart and the other one is the threshold for the λ-GLR chart. The values of such a pair are determined by the specified in-control performance. In this way, development of the p-λ-GLR chart can be straightforward if p-GLR and λ-GLR individual charts are available. The main difficulty in developing a combined chart may

be the control limit determination.

12.6 Derivation of t-GLR chart statistics

In the last section, we proposed a combination method to detect shifts in both p and λ simultaneously. A single chart for the same purpose is provided in this section.

12.6.1 MLE for t-GLR charts

If the Bernoulli parameter and the Poisson mean are both considered, a single GLR control chart for detecting shifts in these two paramters simultaneously is called the t-GLR chart. Based on the log likelihood function for ZIP data in Equation (12.7), the score functions for the t-GLR chart can be presented as

$$\frac{\partial \ln L}{\partial p} = \frac{n_0(e^{-\lambda}-1)}{1-p+pe^{-\lambda}} + \frac{n-n_0}{p}, \tag{12.23}$$

$$\frac{\partial \ln L}{\partial \lambda} = -\frac{n_0 p e^{-\lambda}}{1-p+pe^{-\lambda}} - (n-n_0)\Big(1-\frac{\overline{x}^+}{\lambda}\Big). \tag{12.24}$$

Setting these score functions equal to zeros, we obtain the following score equation system:

$$\frac{n_0(e^{-\lambda}-1)}{1-p+pe^{-\lambda}} + \frac{n-n_0}{p} = 0, \tag{12.25}$$

$$\frac{n_0 p e^{-\lambda}}{1-p+pe^{-\lambda}} + (n-n_0)\Big(1-\frac{\overline{x}^+}{\lambda}\Big) = 0. \tag{12.26}$$

For a given sample, values other than p and λ in the above two equations are constants. To obtain the MLEs for p and λ, i.e., $\hat{p}$ and $\hat{\lambda}$, we need to solve for p and λ, for

which there are several options, such as the Estimation-Maximization method(EM), the elimination method in conjunction with the Newton-Raphson method, and the Newton-Raphson method (NR) for nonlinear equation systems. The most popular form of MLEs used in publications, such as in papers [100, 122, 124], is in the following form

$$\hat{\lambda} = (1 - e^{-\hat{\lambda}})\overline{x}^{+}, \tag{12.27}$$

$$\hat{p} = \frac{n - n_0}{n} \cdot \frac{\overline{x}^{+}}{\hat{\lambda}}. \tag{12.28}$$

The derivation of MLEs based on different methods and more discussion on this topic are presented in Appendix E.

12.6.2 Statistics of t-GLR charts

As long as the MLEs for p and λ can be obtained, one can derive the chart statistics for t-GLR charts as demonstrated in earlier sections for the simpler cases. Suppose the values of p and/or λ have shifted at time point τ^*, then the log likelihood ratio from the changpoint τ^* to the ending time point k can be calculated as

$$\begin{aligned}\beta_{k,\tau}(p_1, \lambda_1) &= \sum_{j=0}^{l} n_j^{k,\tau} \ln \frac{f(j; p_1^{k,\tau}, \lambda_1^{k,\tau})}{f(j; p_0, \lambda_0)} \\ &= n_0^{k,\tau} \ln \frac{1 + (e^{-\lambda_1^{k,\tau}} - 1)p_1^{k,\tau}}{1 + (e^{-\lambda_0} - 1)p_0} + (n^{k,\tau} - n_0^{k,\tau})\Big[\ln \frac{p_1^{k,\tau}}{p_0} + \overline{x}_{k,\tau}^{+} \ln \frac{\lambda_1^{k,\tau}}{\lambda_0} - (\lambda_1^{k,\tau} - \lambda_0)\Big].\end{aligned} \tag{12.29}$$

The statistics of the t-GLR chart can thus be expressed by

$$\begin{aligned} R_k^t &= \max_{0\le\tau<k} \sup_{p_1,\lambda_1} \beta_{k,\tau}(p_1,\lambda_1) \\ &= \max_{0\le\tau<k} \Big\{ n_0^{k,\tau} \ln \frac{1+(e^{-\hat{\lambda}_1^{k,\tau}}-1)\hat{p}_1^{k,\tau}}{1+(e^{-\lambda_0}-1)p_0} + (n^{k,\tau}-n_0^{k,\tau})\Big[\ln\frac{\hat{p}_1^{k,\tau}}{p_0} + \overline{x}_{k,\tau}^{+}\ln\frac{\hat{\lambda}_1^{k,\tau}}{\lambda_0} - (\hat{\lambda}_1^{k,\tau}-\lambda_0)\Big]\Big\}. \end{aligned} \tag{12.30}$$

If the statistics R_k^t are calculated using a moving window with size m, then the statistics will be

$$R_{m,k}^t = \begin{cases} \max\limits_{0\le\tau<k} \Big\{ n_0^{k,\tau} \ln \dfrac{1+(e^{-\hat{\lambda}_1^{k,\tau}}-1)\hat{p}_1^{k,\tau}}{1+(e^{-\lambda_0}-1)p_0} \\ \qquad +(n^{k,\tau}-n_0^{k,\tau})\Big[\ln\frac{\hat{p}_1^{k,\tau}}{p_0} + \overline{x}_{k,\tau}^{+}\ln\frac{\hat{\lambda}_1^{k,\tau}}{\lambda_0} - (\hat{\lambda}_1^{k,\tau}-\lambda_0)\Big]\Big\} & \text{if } k=1,2,\cdots,m \\ \max\limits_{k-m\le\tau<k} \Big\{ n_0^{k,\tau} \ln \dfrac{1+(e^{-\hat{\lambda}_1^{k,\tau}}-1)\hat{p}_1^{k,\tau}}{1+(e^{-\lambda_0}-1)p_0} \\ \qquad +(n^{k,\tau}-n_0^{k,\tau})\Big[\ln\frac{\hat{p}_1^{k,\tau}}{p_0} + \overline{x}_{k,\tau}^{+}\ln\frac{\hat{\lambda}_1^{k,\tau}}{\lambda_0} - (\hat{\lambda}_1^{k,\tau}-\lambda_0)\Big]\Big\} & \text{if } k=m+1,m+2,\cdots. \end{cases} \tag{12.31}$$

Since the t-GLR chart statistic R_k^t is derived under the assumption that p and/or λ may have changed at some time point, R_k^t is a generalization of R_k^p and R_k^λ, which leads to R_k^p and R_k^λ being related to R_k^t. In particular,

1. If $\hat{\lambda}_1 = \lambda_0$, the t-GLR chart statistics in Equation (12.30) will be the p-GLR chart statistics in Equation (12.13);

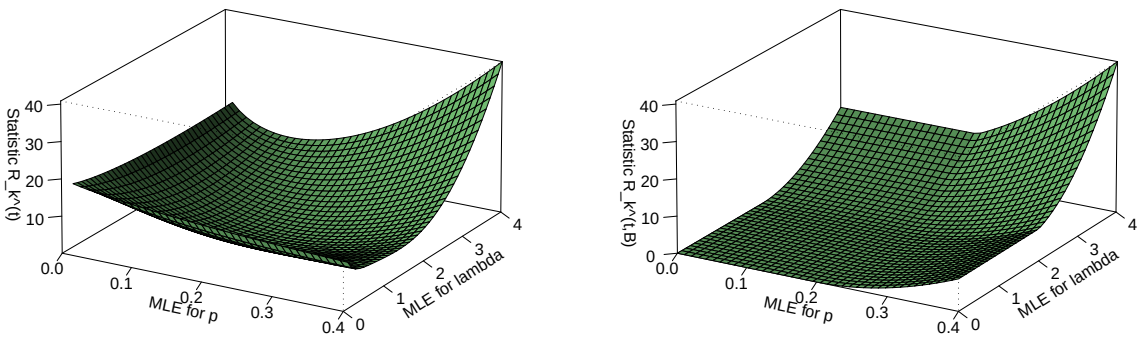
2. If $\hat{p}_1 = p_0$, the t-GLR chart statistics in Equation (12.30) will be the λ-GLR chart statistics in Equation (12.20);

3. If $\hat{p}_1 = p_0 = 1$, the t-GLR chart statistics in Equation (12.30) will be the Poisson-GLR chart statistics in Equation (7.5), which have been developed in Part II.

The statistics R_k^t can be used to detect upward and downward shifts in p and/or λ. To illustrate how the value of R_k^t changes with values of $\hat{p}$ and $\hat{\lambda}$, we can reformat Equations (12.27) and (12.28) to

$$\overline{x}^+ = \frac{\hat{\lambda}}{1 - e^{-\hat{\lambda}}}, \qquad n_0 = n\Big[1 - \hat{p}(1 - e^{-\hat{\lambda}})\Big], \tag{12.32}$$

then we replace $\overline{x}^+$ and n_0 in Equation (12.29) with the above two expressions, leading to the log likelihood ratio $\beta_{k,\tau}(p_1, \lambda_1)$ as a function of $\hat{p}$ and $\hat{\lambda}$ for a given sample. Figure 12.9(a) contains a three dimensional plot of such a function, where $p_0 = 0.2$ and $\lambda_0 = 2$. This curve indicates the minimum value of $\beta_{k,\tau}(p_1, \lambda_1)$ is obtained when $(\hat{p}, \hat{\lambda}) = (0.2, 2)$, and demonstrates that values of $\beta_{k,\tau}(p_1, \lambda_1)$ become larger as points move away from the point $(0.2, 2)$.



(a) **(b)**

Figure 12.9: Relationship between t-GLR chart statistics and maximum likelihood estimates of p and λ, where $p_0 = 0.2$, $\lambda_0 = 2$. (a) MLEs for p and λ; (b) bounded MLEs for p and λ.

One chart can detect any parameter shift in any direction, which is beneficial to users

who are only concerned whether or not a ZIP process is in the in-control condition. On the other hand, if one is interested in a particular combination shift in p and λ, for example an increase in p and a decrease in λ, the general statistic R_k^t may be inefficient. In this sense, more refined statistics for specified shifts are desired. As before a sophisticated sign function can be used to modify R_k^t to detect a pair of specified shifts of interest. However, this modification is more completed than those needed when monitoring single parameters. Another option to the modification is to put suitable bounds on the MLEs. For the specific case where only upward shifts in both p and λ are of interest, the bounded MLEs can be formulated as

$$\hat{p}_1^B = \max(\hat{p}_1, p_0), \qquad \hat{\lambda}_1^B = \max(\hat{\lambda}_1, \lambda_0). \tag{12.33}$$

Use of these bounds can prevent MLEs from ever being less than their in-control values, thus emphasizing the shifts of interest. Figure 12.9 (b) is the corresponding plot of Figure 12.9 (a), with bounds in Equation (12.33) applied. One can see from Figure 12.9 (b) that the curve with $0.2 \leq \hat{p}_1 \leq 0.4$ and $2 \leq \hat{\lambda}_1 \leq 4$ is identical to the corresponding curve in Figure 12.9 (a). In fact, Figure 12.9 (b) can be viewed as an extension of Figure 7.2 from two dimensions to three dimensions. Table 12.6 provides the setup of the bounds for other possible shifts of ZIP parameters.

Table 12.6: Setup of bounds for t-GLR chart MLEs

Shifts of interest		Bounds on MLE	
p	λ	$\hat{p}_1^B$	$\hat{\lambda}_1^B$
upward	upward	$\max(\hat{p}_1, p_0)$	$\max(\hat{\lambda}_1, \lambda_0)$
upward	downward	$\max(\hat{p}_1, p_0)$	$\min(\hat{\lambda}_1, \lambda_0)$
downward	upward	$\min(\hat{p}_1, p_0)$	$\max(\hat{\lambda}_1, \lambda_0)$
downward	downward	$\min(\hat{p}_1, p_0)$	$\min(\hat{\lambda}_1, \lambda_0)$

Chapter 13

Concluding remarks and future work

13.1 Concluding remarks

In contrast to GLR charts for Poisson data, GLR charts for ZIP data becomes more challenging and complicated due to the estimation of an additional parameter and the more extensive computations involved.

Properties of ZIP data appear similar when $\lambda \geq 5$, so use of $\lambda = 2$ and $\lambda = 10$ is recommended for further studies, where $\lambda = 2$ and 10 represent the low and high level of λ, respectively. Because simulations for GLR charts are time-consuming especially when iterative algorithms are used to solve nonlinear equations, the number of simulations is therefore a concern in future studies. The pilot results show that the MLE performance based on 1000 simulations is relatively close to that based on a greater number of simulations.

Solving for the MLE for λ-GLR charts requires numerical methods. Two equations were applied when using the Newton-Raphson method for the solution. The derived GLR statistics based on log likelihood ratios are two-sided. To detect one-sided shifts of interest, we recommended using the sign function applied statistics for both p-GLR

charts and λ-GLR charts. Combining the information of the p-GLR and λ-GLR chart, we conclude that the performance of MLEs for p and λ improves as the value of p increases. The effect of small sample sizes on MLEs is substantial, especially when the values of p and λ are small. The variation of MLEs results in an increase the in-control limits, consequently leading to poor chart performance, i.e., larger detection delay. As reliable parameter estimation methods for ZIP model are not available when sample sizes are small, a simple solution is to ignore cases where sample sizes are less than six, which may potentially postpone the detection by five time periods as compared with an ideal estimation method. However, this easily applied method may be more valuable than the use of the MLEs with small sample sizes producing excessive variation.

To detect shifts in p and λ simultaneously, we can use either a combination method (p-λ-GLR chart), or a single chart (t-GLR chart). The construction of the p-λ-GLR chart is straightforward if p-GLR and λ-GLR charts have been developed. Control limits for p-λ-GLR charts can be obtained using simulation. Solving for MLEs of t-GLR charts also requires some numerical methods. Estimation efficiency of these methods could be a separate study topic. The t-GLR chart uses only one statistic, so this chart is easy to set up in terms of the control limit. A signal produced by the t-GLR chart can indicate a change in the process without knowing the specific parameter shifted. In contrast, one can more easily diagnose the signal given by a p-λ-GLR chart to identify the shifted parameter. The overall performance of p-λ-GLR charts and t-GLR charts can be compared by their detection delay in future studies.

Since the t-GLR chart is designed to detect shifts in either p or λ or both, the statistics of the t-GLR chart can be viewed as a generalization of the statistics for the p-GLR chart, the λ-GLR chart, and the GLR chart for Poisson data. Figure 13.1 illustrates these relationships and summaries our research scheme.

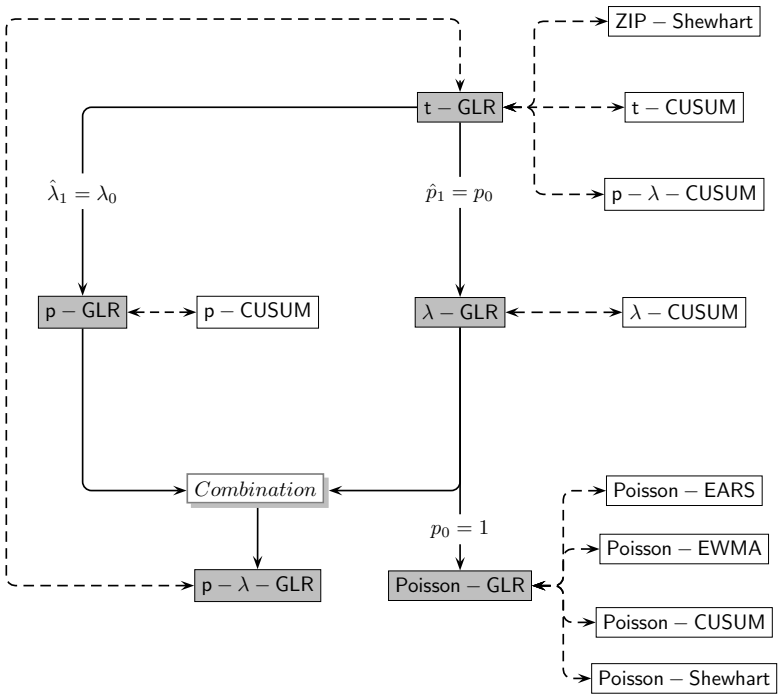


Figure 13.1: Demonstration of research schemes and relationships among GLR charts. (i) Filled rectangles represent GLR charts that we have developed or will be developed; (ii) Blank rectangles represent charts that have been published and can be compared with GLR charts; (iii) Solid lines represent the relationship between the two GLR chart statistics; (iv) Dashed lines represent how two charts can be compared with each other.

13.2 Future work

The ultimate goal of this study is to evaluate the performance of GLR charts for ZIP data for detecting a broad range of parameter shifts. Because time and resources are limited, we have only the derivation and discussion of chart statistics in the dissertation. Future research efforts related to this study could be categorized into the study when a ZIP process is in-control, the study when a ZIP process is out-of-control, and use of GLR charts. They could be briefly summarized as follows:

1. Study when a ZIP process is in-control

   (a) Choosing the window length

   The changepoint framework with the GLR algorithm requires one to calculate the log likelihood ratios for all potential changepoints, which leads to an extensive computational burden. A suitable window size, based on which the GLR statistics are calculated, needs to be chosen before performing the phase II study.

   (b) Determining control limits

   For a given in-control performance, the control limit can be obtained using simulation. To facilitate use of GLR charts, more practical methods for determining the control limits would be provided.

2. Study when a ZIP process is out-of-control

   (a) Comparing GLR charts with other methods

   Each GLR chart would be compared with its alternative methods, and then their overall performance could be evaluated.

   (b) Changing shift types

   GLR chart performance could be evaluated under different types of shift.

3. Application of GLR charts

   (a) Testing Poisson distribution against ZIP distribution

   Because of complexities of the GLR chart for ZIP data, we consider use of a ZIP model only when Poisson models are not suitable. Practical tests for this purpose could be studied in future work.

   (b) Developing an R package for ZIP-GLR charts

   The calculation involved in GLR chart statistics can be complicated due to a number of reasons, including use of the MLE bounds, numerical methods used for solving nonlinear equations, the changepoint scheme, and the window used for relaxing the computational burden. To facilitate use of GLR charts, a bundled R package could be developed.

   (c) Using GLR charts to monitor sparse data in biosurveillance

   GLR charts would be applied to real biosurveillance data, and their performance could be compared with alternative methods.

Appendix A

$R^s_{m,k}$ is an increasing function of $\hat{\lambda}_{1,\hat{\tau}_k,k}$

Given that

$$R_{m,k} = (k - \hat{\tau}_k)\Big[\hat{\lambda}_{1,\hat{\tau}_k,k}\ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big) - (\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0)\Big] \quad \text{and}$$
$$R^s_{m,k} = \operatorname{sgn}(\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0) \cdot R_{m,k},$$

where λ_0, k and $\hat{\tau}_k$ are known positive constants, $k > \hat{\tau}_k$ and $\hat{\lambda}_{1,\hat{\tau}_k,k} > 0$, then $R^s_{m,k}$ is an increasing function of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ over its definitional domain $(0, \infty)$.

Proof. The first derivative of $R_{m,k}$ with respect to $\hat{\lambda}_{1,\hat{\tau}_k,k}$ is

$$\begin{aligned}\frac{dR_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}} &= (k - \hat{\tau}_k)\Big[\ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big) + \hat{\lambda}_{1,\hat{\tau}_k,k}\frac{\lambda_0}{\hat{\lambda}_{1,\hat{\tau}_k,k}}\frac{1}{\lambda_0} - 1\Big] \\ &= (k - \hat{\tau}_k)\ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big).\end{aligned}$$

$R^s_{m,k}$ can be represented as a sub-function as follows:

$$R^s_{m,k} = \text{sgn}(\hat{\lambda}_{1,\hat{\tau}_k,k} - \lambda_0) \cdot R_{m,k} = \begin{cases} R_{m,k} & \text{if} \quad \hat{\lambda}_{1,\hat{\tau}_k,k} > \lambda_0 \\ 0 & \text{if} \quad \hat{\lambda}_{1,\hat{\tau}_k,k} = \lambda_0 \\ -R_{m,k} & \text{if} \quad \hat{\lambda}_{1,\hat{\tau}_k,k} < \lambda_0. \end{cases}$$

Consider the following three cases:

1). If $\hat{\lambda}_{1,\hat{\tau}_k,k} > \lambda_0$, then $R^s_{m,k} = R_{m,k}$, $\frac{dR^s_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}} = \frac{dR_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}} = (k - \hat{\tau}_k)\ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big) > 0$. This indicates that $R^s_{m,k}$ is differentiable over the range (λ_0, ∞), and $R^s_{m,k}$ is an increasing function of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ over that range.

2). If $\hat{\lambda}_{1,\hat{\tau}_k,k} < \lambda_0$, then $R^s_{m,k} = -R_{m,k}$, $\frac{dR^s_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}} = -\frac{dR_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}} = -(k - \hat{\tau}_k)\ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big) > 0$. This indicates that $R^s_{m,k}$ is differentiable over the range $(0, \lambda_0)$, and $R^s_{m,k}$ is an increasing function of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ over that range.

3). If $\hat{\lambda}_{1,\hat{\tau}_k,k} = \lambda_0$, then $R^s_{m,k} = R_{m,k} = 0$. We know that

$$\frac{dR_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}}\bigg|_{\hat{\lambda}_{1,\hat{\tau}_k,k}=\lambda_0} = (k - \hat{\tau}_k)\ln\Big(\frac{\hat{\lambda}_{1,\hat{\tau}_k,k}}{\lambda_0}\Big)\bigg|_{\hat{\lambda}_{1,\hat{\tau}_k,k}=\lambda_0} = 0.$$

To calculate the first derivative of $R^s_{m,k}$ at λ_0, we need to consider the first derivative from the right and left hand sides of λ_0 for the function $R^s_{m,k}$ respectively.

Considering the first derivative from the right hand side, we get

$$\begin{aligned}
\lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^+} \frac{R^s_{m,k}-0}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} &= \lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^+} \frac{R_{m,k}-0}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} \\
&= \lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^+} \frac{R_{m,k}}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} \\
&= \frac{dR_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}}\bigg|_{\hat{\lambda}_{1,\hat{\tau}_k,k}=\lambda_0} \\
&= 0.
\end{aligned}$$

Considering the first derivative from the left hand side, we get

$$\begin{aligned}
\lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^-} \frac{R^s_{m,k}-0}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} &= \lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^-} \frac{-R_{m,k}-0}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} \\
&= \lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^-} -\frac{R_{m,k}}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} \\
&= -\frac{dR_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}}\bigg|_{\hat{\lambda}_{1,\hat{\tau}_k,k}=\lambda_0} \\
&= 0.
\end{aligned}$$

By the definition of the derivative, we have

$$\frac{dR^s_{m,k}}{d\hat{\lambda}_{1,\hat{\tau}_k,k}}\bigg|_{\hat{\lambda}_{1,\hat{\tau}_k,k}=\lambda_0} = \lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^+} \frac{R^s_{m,k}-0}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} = \lim_{\hat{\lambda}_{1,\hat{\tau}_k,k}\to\lambda_0^-} \frac{R^s_{m,k}-0}{\hat{\lambda}_{1,\hat{\tau}_k,k}-\lambda_0} = 0.$$

Combining the above three cases, $R^s_{m,k}$ is differentiable over its definitional domain $(0,\infty)$, and its first derivative is greater than zero anywhere except at the point $\hat{\lambda}_{1,\hat{\tau}_k,k}=\lambda_0$, where its first derivative is zero. The statistic $R^s_{m,k}$ is consequently an increasing function of $\hat{\lambda}_{1,\hat{\tau}_k,k}$ over its definitional domain $(0,\infty)$. □

Appendix B

CDF Relationship between statistics of Shewhart and GLR with $m = 1$

Suppose the GLR chart with $m = 1$, then $\tau = k - 1$, $\hat{\lambda}_{1,\tau,k} = x_k$ in Equation (7.13) and

$$R^s_{m,k} = \left[x_k \ln(\frac{x_k}{\lambda_0}) - (x_k - \lambda_0) \right] \cdot \text{sgn}(x_k - \lambda_0), \tag{B.1}$$

where X_k is a Poisson random variable with mean λ, i.e., $X_k \sim \text{Poisson}(\lambda)$. Let $F_X(x_k)$ be the cumulative density function (cdf) of x_k .

For simplicity, we can write Equation (B.1) in the following form

$$R = R^s_{m,k} = f(x_k). \tag{B.2}$$

That means the random variable R is a transformation of the random variable x_k. By the statement in Appendix A, $f(x_k)$ is an increasing function of x_k on its support $\{0, 1, 2, \cdots\}$. Thus, $r = f(x_k)$ has an inverse function, denoted by $f^{-1}(r)$.

The statistic of the Shewhart chart is

$$S_k = \frac{x_k - \lambda_0}{\sqrt{\lambda_0}}. \tag{B.3}$$

For simplicity, we can write Equation (B.3) in the following form

$$S = g(x_k). \tag{B.4}$$

Equation (B.3) means the random variable S is another transformation of x_k. It is obvious that $g(x_k)$ is an increasing function of x_k on its support $\{0, 1, 2, \cdots\}$. Thus, $s = g(x_k)$ has an inverse function, denoted by $g^{-1}(s)$.

Suppose $F_R(r)$ and $F_S(s)$ are the cdfs of the random variable R and S, respectively, then we have

$$F_R(r) = F_X[f^{-1}(r)] \tag{B.5}$$

$$F_S(s) = F_X[g^{-1}(s)]. \tag{B.6}$$

Thus, for a given value of x_k, we obtain the cdf values of R and S as follows:

$$F_R(r) = F_R[f(x_k)] = F_X[f^{-1}(f(x_k))] = F_X(x_k) \tag{B.7}$$

$$F_S(s) = F_S[g(x_k)] = F_X[g^{-1}(g(x_k))] = F_X(x_k). \tag{B.8}$$

Equations (B.7) and (B.8) imply that, for any given $x_k \in \{0, 1, 2, \cdots\}$, the random variable R and S have the same cdf value as the random variable X_k.

Appendix C

More control limits for GLR, EARS, Shewhart, CUSUM and EWMA methods

Table C.1: More control limits for GLR, EARS, Shewhart, CUSUM and EWMA charts for Poisson data with $\lambda_0 = 2$.

		GLR	EARS	Shewhart	CUSUM							EWMA				
ATS_0	FAR	—	—	—	0.25	0.5	1	1.5	2	2.5	3	0.05	0.1	0.2	0.3	0.4
50	0.02000	2.5903	3.0230	2.1215	6.3178	5.3002	3.7149	3.1317	2.5846	2.0543	1.5487	2.3265	2.5680	2.9613	3.3126	3.6016
75	0.01333	2.5998	3.4199	2.8282	7.7617	6.2901	4.4852	3.3316	2.7934	2.5278	2.2718	2.3814	2.6485	3.0708	3.4564	3.7751
100	0.01000	3.1639	3.5716	2.8285	8.7904	6.9768	5.0664	4.1313	3.3731	2.5827	2.2740	2.4193	2.7021	3.1484	3.5287	3.9182
150	0.00667	3.7646	3.8578	2.8291	10.4204	7.9801	5.7133	4.2649	3.7910	3.1616	2.5497	2.4704	2.7735	3.2543	3.6667	4.0147
200	0.00500	3.8690	4.0309	2.8297	11.5957	8.9143	6.1331	5.1300	3.9512	3.5281	3.2710	2.5051	2.8212	3.3232	3.7582	4.1424
250	0.00400	4.1251	4.2840	3.5341	12.6207	9.3233	6.5559	5.1978	4.3730	3.5821	3.2729	2.5331	2.8575	3.3737	3.8166	4.2386
300	0.00333	4.4178	4.4281	3.5342	13.4940	9.9723	6.8441	5.2645	4.5817	4.0506	3.2749	2.5537	2.8860	3.4117	3.8632	4.3110
350	0.00286	4.7416	4.5427	3.5344	14.2429	10.3219	7.1336	5.3978	4.5844	4.0528	3.5464	2.5710	2.9097	3.4476	3.9128	4.3671
400	0.00250	4.8582	4.5710	3.5345	14.8711	10.9426	7.4238	5.5936	4.7916	4.1050	3.5481	2.5864	2.9302	3.4786	3.9418	4.3988
450	0.00222	5.0924	4.7143	3.5347	15.4380	11.2793	7.6853	6.1324	4.7944	4.5253	3.5499	2.5989	2.9474	3.5024	3.9769	4.4122
500	0.00200	5.0964	4.7160	3.5348	15.9979	11.6138	7.8404	6.1973	5.1605	4.5271	3.8177	2.6100	2.9635	3.5247	4.0084	4.4581
600	0.00167	5.1838	4.8568	3.5351	16.9244	11.9847	8.1337	6.2649	5.3720	4.5806	4.2704	2.6300	2.9892	3.5624	4.0575	4.5093
700	0.00143	5.3449	4.9985	3.5354	17.7177	12.6102	8.4250	6.3980	5.5803	4.5841	4.2721	2.6461	3.0117	3.5945	4.0987	4.5594
800	0.00125	5.4362	5.0527	3.5357	18.3845	12.9520	8.7110	6.5291	5.5828	5.0504	4.2738	2.6591	3.0296	3.6185	4.1282	4.6243
900	0.00111	5.6228	5.1428	3.5359	19.0175	13.2944	8.8443	7.0672	5.7463	5.0521	4.5450	2.6702	3.0467	3.6433	4.1567	4.6427
1000	0.00100	5.6962	5.1972	4.2420	19.6040	13.6282	9.1147	7.1970	5.7928	5.0538	4.5461	2.6810	3.0613	3.6648	4.1842	4.6898
1100	0.00091	5.8722	5.2852	4.2421	20.1523	13.9533	9.2044	7.2616	5.9522	5.1060	4.5472	2.6901	3.0745	3.6838	4.2130	4.7212
1200	0.00083	6.1180	5.4282	4.2422	20.6322	14.2760	9.4230	7.3303	6.1609	5.1093	4.5483	2.6978	3.0856	3.7021	4.2285	4.7522
1300	0.00077	6.2593	5.4291	4.2422	21.0555	14.5676	9.5555	7.3338	6.1636	5.5256	4.5494	2.7063	3.0960	3.7183	4.2499	4.7849
1400	0.00071	6.3232	5.4782	4.2423	21.4894	14.6331	9.7113	7.3975	6.3704	5.5277	4.8154	2.7130	3.1057	3.7309	4.2710	4.7986
1500	0.00067	6.3259	5.5702	4.2424	21.8631	14.9428	9.8407	7.4619	6.3720	5.5297	4.8169	2.7192	3.1149	3.7438	4.2904	4.7996
2000	0.00050	6.5357	5.7147	4.2427	23.4611	15.9082	10.3344	8.1344	6.5830	5.6376	5.2709	2.7452	3.1516	3.7987	4.3637	4.8792
2500	0.00040	6.7081	5.8564	4.2430	24.7183	16.5909	10.7778	8.3320	6.9505	6.0527	5.2748	2.7656	3.1810	3.8403	4.4172	4.9575
3000	0.00033	7.0120	5.9984	4.2433	25.8056	17.2322	11.1321	8.4620	7.1623	6.1069	5.5464	2.7818	3.2034	3.8730	4.4571	5.0173

Appendix D

MLE for λ-GLR charts

As discussed in Section 12.4.1, the MLE for λ of λ-GLR charts can be expressed as

$$\frac{p_0}{p_0 + (1 - p_0)e^{\hat{\lambda}}} = \frac{n - n_0}{n_0}\Big(\frac{\overline{x}^+}{\hat{\lambda}} - 1\Big), \tag{D.1}$$

where n, n_0, p_0, and $\overline{x}^+$ are constants in a given sample.

One of popular numerical methods for solving this type of equation is the Newton-Raphson algorithm. It is known that the representation of equations may influence the feasibility or efficiency of the Newton-Raphson method. In other words, the Newton-Raphson method may not work for an expression, but may work for its alternatives. Some of the variables involved in Equation (D.1) are random in practice and in any simulation study. It is therefore desired to find a suitable expression of Equation (D.1), which makes the Newton-Raphson algorithm work for all possible parameter and variable combinations. Unfortunately, preliminary simulation indicates that it is hard to find a single expression that works for various settings. On the other hand, some combinations of alternative equations may fulfill our purpose. One of the functional combinations can be

expressed as

$$f_1(\lambda) = np_0\lambda - (n - n_0)\left[p_0\overline{x}^+ + (1 - p_0)(\overline{x}^+ - \lambda)e^{\lambda}\right], \tag{D.2}$$

$$f_2(\lambda) = \lambda - (1 - e^{-\lambda})\overline{x}^+. \tag{D.3}$$

It can be seen that Equation (D.3) is not an equivalent format of Equation (D.1); it is actually an equation used for the MLE for t-GLR charts, which will be shown in Appendix E. Equation (D.3) is used only when $\overline{x}^+ > 15$.

As indicated in the following Newton-Raphson iteration,

$$\lambda_{i+1} = \lambda_i - \frac{f(\lambda_i)}{f'(\lambda_i)}, \quad i = 0, 1, 2, \cdots, \tag{D.4}$$

the first derivative with respect to λ is required. The corresponding derivatives of Equations (D.2) and (D.3) are given by

$$f_1'(\lambda) = np_0 - (n - n_0)(1 - p_0)(\overline{x}^+ - \lambda - 1)e^{\lambda}, \tag{D.5}$$

$$f_2'(\lambda) = 1 - \overline{x}^+ e^{-\lambda}. \tag{D.6}$$

The R code for solving for $\hat{\lambda}$ can be developed as follows. The main function is used to implement the iteration expressed in Equation (D.4), and two separate functions are used to solve for $\hat{\lambda}$ in Equations (D.2) and (D.3).

```
#=== Main function: Newton-Raphson iteration ===
 newtonraphson <- function(ftn.fx, x0, tol = 1e-6, max.iter = 100)
{
  # 1. The Newton-Raphson algorithm for solving ftn.fx(x)[1] = 0;
  # 2. We assume that ftn.fx(x) is a function of a single variable
```

```
  #    that returns the function value and the first derivative as
  #    a vector of length 2;
  # 3. x0 is the initial guess at the root;
  # 4. The algorithm terminates when the function value is within
  #    distance tol of 0, or the number of iterations exceeds
  #    max.iter value.

  # Initialize
    x <- x0
    iter <- 0
    fx <- ftn.fx(x)

  # Continue iterating until stopping conditions are met
    while ( (abs(fx[1]) > tol) && (iter < max.iter) ) {
       x <- x - fx[1]/fx[2]
       fx <- ftn.fx(x)
       iter <- iter + 1
       cat( "At iteration", iter, "value of x is:", x, "\n" )
    }

  # Output depends on success of the algorithm
    if ( abs(fx[1]) > tol ) {
       cat( "Algorithm failed to converge\n" )
       return(NULL)
    } else {
       cat("Algorithm converged\n")
       return(x)
    }
}

#=== Solve Expression 1 ===
ftn.fx1 <- function(x) {
  # 1. p0 is the value of the Bernoulli parameter
       when in-control;
  # 2. n is the sample size;
  # 3. n0 is the number of zeros in a sample;
  # 4. xbar is the mean of positive counts in a sample.

    term1 <- n - n0
    term2 <- (xbar - x)*exp(x)
    term3 <- n*p0*x
    fx <-  term3 - term1*( p0*xbar + (1 - p0)*term2 )
    # fx is the expression for solving
```

```
    dfx <- n*p0 - (n - n0)*(1 - p0)*(xbar - x - 1)*exp(x)
    # dfx is the first derivative with respect to lambda
    return( c(fx, dfx) )
    # Return vector (fx, dfx)
}

#=== Solve Expression 2 ===
ftn.fx2 <- function(x) {
     term1 <- n - n0
     term2 <- (xbar - x)*exp(x)
     term3 <- n*p*x
     fx <-  term3 - term1*( p*xbar + (1 - p)*term2 )
     # fx is the expression for solving
     dfx <- n*p -(n - n0)*(1 - p)*(xbar - x - 1)*exp(x)
     # dfx is the first derivative with respect to lambda
     return( c(fx, dfx) )
     # Return vector (fx, dfx)
}
```

Appendix E

MLEs for t-GLR charts

As shown in Section 12.6.1, the score equations for p and λ can be written as

$$\frac{n_0(e^{-\lambda}-1)}{1-p+pe^{-\lambda}}+\frac{n-n_0}{p}=0, \tag{E.1}$$

$$\frac{n_0 pe^{-\lambda}}{1-p+pe^{-\lambda}}+(n-n_0)\left(1-\frac{\overline{x}^{+}}{\lambda}\right)=0. \tag{E.2}$$

where the definitions of p, λ, n, n_0 and $\overline{x}^{+}$ are the same as those given in Equation (12.7).

The MLEs for p and λ of the t-GLR chart are the roots of this equation system. There are several options to solve this system, such as the expectation-maximization (EM) method [100, 105, 110, 117], the elimination method in conjunction with the Newton-Raphson method [100, 122, 124], and the Newton-Raphson method for nonlinear equation systems. The first two methods work in two steps: (1) estimate one parameter, (2) calculate the other parameter estimate based on the first step. In contrast, the Newton-Raphson method for nonlinear equation systems solves for these two parameters simultaneously, which may produce more accurate results. An evaluation on MLE performance of these algorithms may give rise to a separate research topic. In the following sections, application of the elimination method and the Newton-Raphson method will be demonstrated.

E.1 Elimination method

Reformatting Equation (E.1), we can obtain

$$p = \frac{n - n_0}{n} \cdot \frac{1}{1 - e^{-\lambda}}. \tag{E.3}$$

Note that p in Equation (E.3) is derived only based on the score function of p, so the MLE for p for p-GLR charts is in the same form.

Replacing p in Equation (E.2) with Equation (E.3) and making some simplifications, we obtain

$$\lambda = (1 - e^{-\lambda})\overline{x}^{+}. \tag{E.4}$$

Using Equation (E.4), Equation (E.3) can be reduced to

$$p = \frac{n - n_0}{n} \cdot \frac{\overline{x}^{+}}{\lambda}. \tag{E.5}$$

So the MLEs of p and λ for t-GLR charts can be calculated as follows:

$$\hat{\lambda} = (1 - e^{-\hat{\lambda}})\overline{x}^{+}, \tag{E.6}$$

$$\hat{p} = \frac{n - n_0}{n} \cdot \frac{\overline{x}^{+}}{\hat{\lambda}}. \tag{E.7}$$

The root $\hat{\lambda}$ of Equation (E.6) can be solved using the algorithm given in Appendix D, and then the value of $\hat{p}$ in Equation (E.7) can be computed.

E.2 Newton-Raphson method for nonlinear equation systems

Score equations for p and λ in Equations (E.1) and (E.2) are nonlinear. In order to obtain $\hat{p}$ and $\hat{\lambda}$, we can use the Newton-Raphson method which is developed to solve nonlinear equation systems. To apply this method easily, we can rewrite Equations (E.1) and (E.2) in the following equivalent expressions

$$S_1(\beta) = n - n_0 + np(e^{-\lambda} - 1), \tag{E.8}$$

$$S_2(\beta) = np\lambda - (n - n_0)\left[p\overline{x}^+ + (1-p)(\overline{x}^+ - \lambda)e^{\lambda}\right], \tag{E.9}$$

and introduce new notation

$$\beta = \begin{pmatrix} p \\ \lambda \end{pmatrix}, \qquad \hat{\beta} = \begin{pmatrix} \hat{p} \\ \hat{\lambda} \end{pmatrix}, \qquad S(\beta) = \begin{pmatrix} S_1(\beta) \\ S_2(\beta) \end{pmatrix}.$$

If we let $D(\beta)$ denote the 2×2 matrix for partial derivatives of $S(\beta)$,

$$D(\beta) = \begin{pmatrix} \frac{\partial S_1(\beta)}{\partial p} & \frac{\partial S_1(\beta)}{\partial \lambda} \\ \frac{\partial S_2(\beta)}{\partial p} & \frac{\partial S_2(\beta)}{\partial \lambda} \end{pmatrix},$$

where

$$\begin{aligned}
\frac{\partial S_1(\beta)}{\partial p} &= n(e^{-\lambda} - 1), \\
\frac{\partial S_1(\beta)}{\partial \lambda} &= -npe^{-\lambda}, \\
\frac{\partial S_2(\beta)}{\partial p} &= n\lambda + (n - n_0)\left[(\overline{x}^+ - \lambda)e^{\lambda} - \overline{x}^+\right], \\
\frac{\partial S_2(\beta)}{\partial \lambda} &= np - (n - n_0)(1 - p)(\overline{x}^+ - \lambda - 1)e^{\lambda},
\end{aligned}$$

then using the first order Taylor series expansion of $S(\hat{\beta})$ about β and setting the expansion equal to the 2×1 vector of zeros gives

$$\hat{\beta} = \beta + D(\beta)^{-1}S(\beta) = \beta + \gamma(\beta). \tag{E.10}$$

This expression gives rise to the following Newton-Raphson algorithm.

1. Obtain suitable initial values for β as $\beta_0 = \begin{pmatrix} p_0 \\ \lambda_0 \end{pmatrix} = \begin{pmatrix} \frac{n-n_0}{n} \\ \overline{x}^+ \end{pmatrix}$;
2. Compute $D(\beta_0)$ and $S(\beta_0)$ and compute $\gamma(\beta_0)$;
3. Compute $\hat{\beta}_1 = \beta_0 + \gamma(\beta_0)$;
4. If $\hat{\beta}_1 \approx \beta_0$, then declare $\hat{\beta} = \beta_1$. Otherwise, replace β_0 by $\hat{\beta}_1$ and return to step 2.

Bibliography

[1] R.D. Fricker Jr, B.L. Hegler, and D.A. Dunfee. Comparing syndromic surveillance detection methods: EARS' versus a CUSUM-based methodology. *Statistics in Medicine*, 27(17):3407–3429, 2008.

[2] D.L. Buckeridge, H. Burkom, M. Campbell, W.R. Hogan, and A.W. Moore. Algorithms for rapid outbreak detection: a research synthesis. *Journal of Biomedical Informatics*, 38(2):99–113, 2005.

[3] J.I. Tokars, H. Burkom, J. Xing, R. English, S. Bloom, K. Cox, and J.A. Pavlin. Enhancing time-series detection algorithms for automated biosurveillance. *Emerging Infectious Diseases*, 15(4):533–539, 2009.

[4] W.H. Woodall. The use of control charts in health-care and public-health surveillance (with discussion). *Journal of Quality Technology*, 38(2):89–104, 2006.

[5] T. Lotze, S. Murphy, and G. Shmueli. Implementation and comparison of preprocessing methods for biosurveillance data. *Advances in Disease Surveillance*, 6: 1–20, 2008.

[6] W.H. Woodall, J.B. Marshall, M.D. Joner Jr, S.E. Fraker, and A.S.G. Abdel-Salam. On the use and evaluation of prospective scan methods for health-related surveillance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(1):223–237, 2008.

[7] L. Hutwagner, W. Thompson, G.M. Seeman, and T. Treadwell. The bioterrorism preparedness and response early aberration reporting system (EARS). *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 80(Supplement 1):89–96, 2003.

[8] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel. *Time series analysis: Forecasting and control*. Oakland CA: Holden-Day, 1976.

[9] R.H. Shumway and D.S. Stoffer. *Time series analysis and its applications*. Springer Verlag, 2000.

[10] G. Shmueli and H. Burkom. Statistical challenges facing early outbreak detection in biosurveillance. *Technometrics*, 52(1):39–51, 2010.

[11] A. R. McLean and R. M. Anderson. Measles in developing countries part I. epidemiological parameters and patterns. *Epidemiology and Infection*, 100(1):111–133, 1988.

[12] Centers for disease control and prevention. `http://www.bt.cdc.gov//surveillance/ears`, Accessed February 20, 2010.

[13] Centers for disease control and prevention. real-time hospital data user guide: Application version 2.11. `http://www.cdc.gov/BioSense/files/CDC_BioSense_BioSense_Hospital_Data_User_Guide_V2.11.pdf`, Accessed February 20, 2010.

[14] W.S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836, 1979.

[15] G.D. Williamson and G.W. Hudson. A monitoring system for detecting aberrations in public health surveillance reports. *Statistics in Medicine*, 18(23):3283–3298, 1999.

[16] D.C. Montgomery, C.L. Jennings, and M. Kulahci. *Introduction to time series analysis and forecasting*. John Wiley, 2008.

[17] D.L. Buckeridge, A. Okhmatovskaia, S. Tu, M. O'Connor, C. Nyulas, and M.A. Musen. Understanding detection performance in public health surveillance: Modeling aberrancy-detection algorithms. *Journal of the American Medical Informatics Association*, 15(6):760–769, 2008.

[18] H.S. Burkom, S.P. Murphy, and G. Shmueli. Automated time series forecasting for biosurveillance. *Statistics in Medicine*, 26(22):4202–4218, 2007.

[19] T.H. Lotze and G. Shmueli. How does improved forecasting benefit detection? an application to biosurveillance. *International Journal of Forecasting*, 25(3):467 – 483, 2009.

[20] Centers for disease control and prevention. `http://www.bt.cdc.gov/surveillance/ears/datasets.asp`, Accessed February 20, 2010.

[21] L. Hutwagner, T. Browne, G.M. Seeman, and A.T. Fleischauer. Comparing aberration detection methods with simulated data. *Emerging Infectious Disease*, 11(2): 314–316, 2005.

[22] J.L. Szarka III, L. Gan, and W.H. Woodall. Comparison of the early aberration reporting system (EARS) W2 methods to an adaptive threshold method. *Statistics in Medicine*, 30(5):489–504, 2011.

[23] S.E. Fraker. Performance metrics for surveillance schemes. *Quality Engineering*, 20 (4):451–464, 2008.

[24] P.A. Rogerson and I. Yamada. Approaches to syndromic surveillance when data consist of small regional counts. *Morbidity and Mortality Weekly Report*, 54 (Supplement):79–85, 2004.

[25] M. Kulldorff, Z. Zhang, J. Hartman, R. Heffernan, L. Huang, and F. Mostashari. Benchmark data and power calculations for evaluating disease outbreak detection methods. *Morbidity and Mortality Weekly Report*, 54(Supplement):144–151, 2004.

[26] M. Frisén. Optimal sequential surveillance for finance, public health, and other areas. *Sequential Analysis*, 28(3):310–337, 2009.

[27] J.C. Benneyan, R.C. Lloyd, and P.E. Plsek. Statistical process control as a tool for research and healthcare improvement. *Quality & Safety in Health Care*, 12(6):458–464, 2003.

[28] W. A. Shewhart. *Economic control of quality of manufactured product*. Van Nostrand, New York, 1931.

[29] E.S. Page. Continuous inspection schemes. *Biometrika*, 41(1-2):100–115, 1954.

[30] S.W. Roberts. Control chart tests based on geometric moving averages. *Technometrics*, 1:239–250, 1959.

[31] J.M. Lucas and R.B. Crosier. Fast initial response for CUSUM quality-control schemes: give your CUSUM a head start. *Technometrics*, 24(3):199–205, 1982.

[32] Z. Wu, M. Yang, W. Jiang, and M.B.C. Khoo. Optimization designs of the combined Shewhart-CUSUM control charts. *Computational Statistics & Data Analysis*, 53(2): 496–506, 2008.

[33] Y. Zhao, F. Tsung, and Z. Wang. Dual CUSUM control schemes for detecting a range of mean shifts. *IIE Transactions*, 37(11):1047–1057, 2005.

[34] R.S. Sparks. CUSUM charts for signalling varying location shifts. *Journal of Quality Technology*, 32(2):157–171, 2000.

[35] W. Jiang, L. Shu, and D.W. Apley. Adaptive CUSUM procedures with EWMA-based shift estimators. *IIE Transactions*, 40(10):992–1003, 2008.

[36] L. Shu, W. Jiang, and Z. Wu. Adaptive CUSUM procedures with Markovian mean estimation. *Computational Statistics & Data Analysis*, 52(9):4395–4409, 2008.

[37] A.S. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica*, 12(6):601–611, 1976.

[38] D. Siegmund and ES Venkatraman. Using the generalized likelihood ratio statistic for sequential detection of a change-point. *The Annals of Statistics*, 23(1):255–271, 1995.

[39] T.L. Lai. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 57(4): 613–658, 1995.

[40] T.L. Lai. Sequential analysis: some classical problems and new challenges. *Statistica Sinica*, 11(2):303–350, 2001.

[41] D.W. Apley and J. Shi. The GLRT for statistical process control of autocorrelated processes. *IIE Transactions*, 31(12):1123–1134, 1999.

[42] G. Capizzi. Design of change detection algorithms based on the generalized likelihood ratio test. *Environmetrics*, 12(8):749–756, 2001.

[43] D.M. Hawkins, P. Qiu, and W.K. Chang. The changepoint model for statistical process control. *Journal of Quality Technology*, 35(4):355–366, 2003.

[44] D.M. Hawkins and K.D. Zamba. Statistical process control for shifts in mean or variance using a changepoint formulation. *Technometrics*, 47(2):164–173, 2005.

[45] D. Han and F. Tsung. Comparison of the cuscore, GLRT and cusum control charts for detecting a dynamic mean change. *Annals of the Institute of Statistical Mathematics*, 57(3):531–552, 2005.

[46] M. Höhle and M. Paul. Count data regression charts for the monitoring of surveillance time series. *Computational Statistics & Data Analysis*, 52(9):4357–4368, 2008.

[47] M.R. Reynolds Jr and J. Lou. An evaluation of a GLR control chart for monitoring the process mean. *Journal of Quality Technology*, 42(3):287–310, 2010.

[48] W. Huang. GLR control chart for monitoring a proportion. *Dissertation proposal research*, 2010.

[49] K.L. Tsui, W. Chiu, P. Gierlich, D. Goldsman, X. Liu, and T. Maschek. A review of healthcare, public health, and syndromic surveillance. *Quality Engineering*, 20(4): 435–450, 2008.

[50] A.L. Lloyd. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1470):985–993, 2001.

[51] M.R. Reynolds Jr and Z.G. Stoumbos. Individuals control schemes for monitoring the mean and variance of processes subject to drifts. *Stochastic Analysis and Applications*, 19(5):863–892, 2001.

[52] S. Gubbins, S. Carpenter, M. Baylis, J.L.N. Wood, and P.S. Mellor. Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number. *Journal of the Royal Society Interface*, 5(20):363–371, 2008.

[53] H.S. Burkom, L. Hutwagner, and R. Rodriguez. Using point-source epidemic curves to evaluate alerting algorithms for biosurveillance. *2004 Proceedings of the American Statistical Association, Statistics in Government Section [CD-ROM], Toronto*, 2005.

[54] D. Han, F. Tsung, X. Hu, and K. Wang. CUSUM and EWMA multi-charts for detecting a range of mean shifts. *Statistica Sinica*, 17(3):1139–1164, 2007.

[55] S.W. Han, K.L. Tsui, B. Ariyajunya, and S.B. Kim. A comparison of CUSUM, EWMA, and temporal scan statistics for detection of increases in Poisson rates. *Quality and Reliability Engineering International*, 26(3):279–289, 2010.

[56] M.C. Testik, BD McCullough, and C.M. Borror. The effect of estimated parameters on Poisson EWMA control charts. *Quality Technology and Quantitative Management*, 3:513–527, 2006.

[57] C.H. Weiß. Controlling correlated processes of Poisson counts. *Quality and Reliability Engineering International*, 23(6):741–754, 2007.

[58] M. Stoto, R. Fricker, A. Jain, A. Diamond, J. Davies-Cole, C. Glymph, G. Kidane, G. Lum, L.V. Jones, and K. Dehan. *Evaluating statistical methods for syndromic surveillance*. Springer, 2006.

[59] W.A. Shewhart. *Statistical method from the viewpoint of quality control*. Dover Publications, 1986.

[60] C.W. Champ and W.H. Woodall. Exact results for Shewhart control charts with supplementary runs rules. *Technometrics*, 29(4):393–399, 1987.

[61] D.C. Montgomery. *Introduction to statistical quality control*. John Wiley & Sons, 2007.

[62] W.H. Woodall. Control charts based on attribute data: bibliography and review. *Journal of Quality Technology*, 29(2):172–183, 1997.

[63] D. Brook and DA Evans. An approach to the probability distribution of CUSUM run length. *Biometrika*, 59(3):539, 1972.

[64] J.M. Lucas. Counted data CUSUM's. *Technometrics*, 27(2):129–144, 1985.

[65] D.M. Hawkins and D.H. Olwell. *Cumulative sum charts and charting for quality improvement*. Springer Verlag, 1998.

[66] G.B. Hill, C.C. Spicer, and J.A.C. Weatherall. The computer surveillance of congenital malformations. *British Medical Bulletin*, 24(3):215, 1968.

[67] J.A.C. Weatherall and J.C. Haskey. Surveillance of malformations. *British Medical Bulletin*, 32(1):39–44, 1976.

[68] L.C. Hutwagner, E.K. Maloney, N.H. Bean, L. Slutsker, and S.M. Martin. Using laboratory-based surveillance data for prevention: An algorithm for detecting Salmonella outbreaks. *Emerging Infectious Diseases*, 3(3):395–400, 1997.

[69] A.G. Ryan and W.H. Woodall. Control charts for Poisson count data with varying sample sizes. *Journal of Quality Technology*, 42(3):260–275, 2010.

[70] S.V. Crowder. Computation of ARL for combined individual measurement and moving range charts. *Journal of Quality Technology*, 19(2):98–102, 1987.

[71] S.V. Crowder. A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics*, 29(4):401–407, 1987.

[72] J.M. Lucas and M.S. Saccucci. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, 32(1):1–12, 1990.

[73] E. Yashchin. Statistical control schemes: methods, applications and generalizations. *International Statistical Review/Revue Internationale de Statistique*, 61(1):41–66, 1993.

[74] C.M. Borror, C.W. Champ, and S.E. Rigdon. Poisson EWMA control charts. *Journal of Quality Technology*, 30(4):352–361, 1998.

[75] S.V. Crowder and M.D. Hamilton. An EWMA for monitoring a process standard deviation. *Journal of Quality Technology*, 24(1):12–21, 1992.

[76] W.H. Woodall and M.A. Mahmoud. The inertial properties of quality control charts. *Technometrics*, 47(4):425–436, 2005.

[77] R.D. Fricker Jr. Some methodological issues in biosurveillance. *Statistics in Medicine*, 30(5):403–415, 2011.

[78] G. Lorden. Procedures for reacting to a change in distribution. *The Annals of Mathematical Statistics*, 42(6):1897–1908, 1971.

[79] M. Basseville and I.V. Nikiforov. *Detection of abrupt changes: theory and application*. Prentice-Hall, 1993.

[80] M. Basseville and A. Benveniste. Desgin and comparative study of some sequential jump detection algorithms for digital signals. *Speech and Signal Processing, IEEE Transactions on Acoustics*, 31(3):521–535, 1983.

[81] D.V. Hinkley. Inference about the change-point in a sequence of random variables. *Biometrika*, 57(1):1, 1970.

[82] G.C. Runger and M.C. Testik. Control charts for monitoring fault signatures: Cuscore versus GLR. *Quality and Reliability Engineering International*, 19(4):387–396, 2003.

[83] S.M. Kay and J.R. Gabriel. An invariance property of the generalized likelihood ratio test. *Signal Processing Letters, IEEE*, 10(12):352–355, 2003.

[84] A. Chen and Y.K. Chen. Design of EWMA and CUSUM control charts subject to random shift sizes and quality impacts. *IIE Transactions*, 39(12):1127–1141, 2007.

[85] J.H. Ryu, H. Wan, and S. Kim. Optimal design of a CUSUM chart for a mean shift of unknown size. *Journal of Quality Technology*, 42(3), 2010.

[86] M.R. Reynolds Jr and Z.G. Stoumbos. Control charts and the efficient allocation of sampling resources. *Technometrics*, 46(2):200–214, 2004.

[87] R. Domangue and S.C. Patch. Some omnibus exponentially weighted moving average statistical process monitoring schemes. *Technometrics*, 33(3):299–313, 1991.

[88] J. Stoer and R. Bulirsch. *Introduction to numerical analysis*. Springer Verlag, 2002.

[89] T. Fawcett and F. Provost. Activity monitoring: Noticing interesting changes in behavior. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 53–62. ACM, 1999.

[90] Y. Shen, W.K. Wong, and G.F. Cooper. A generalization of the AMOC curve. *Advances in Disease Surveillance*, 1(1):65, 2006.

[91] X. Jiang, G.F. Cooper, and D.B. Neill. Generalized AMOC curves for evaluation and improvement of event surveillance. In *AMIA Annual Symposium Proceedings*, pages 281–285. American Medical Informatics Association, 2009.

[92] G.F. Cooper, D.H. Dash, J.D. Levander, W.K. Wong, W.R. Hogan, and M.M. Wagner. Bayesian biosurveillance of disease outbreaks. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pages 94–103. AUAI Press, 2004.

[93] G.F. Cooper, J.N. Dowling, J.D. Levander, and P. Sutovsky. A Bayesian algorithm for detecting CDC category A outbreak diseases from emergency department chief complaints. *Advances in Disease Surveillance*, 2(2):45, 2007.

[94] L. Shu and W. Jiang. A Markov chain model for the adaptive CUSUM control chart. *Journal of Quality Technology*, 38(2):135–147, 2006.

[95] T.L. Lai. Information bounds and quick detection of parameter changes in stochastic systems. *IEEE Transactions on Information Theory*, 44(7):2917–2929, 1998.

[96] M.R. Reynolds Jr and Z.G. Stoumbos. Comparisons of some exponentially weighted moving average control charts for monitoring the process mean and variance. *Technometrics*, 48(4):550–567, 2006.

[97] B.Y. Reis and K.D. Mandl. Time series modeling for syndromic surveillance. *BMC Medical Informatics and Decision Making*, 3(1):2, 2003.

[98] J.C. Brillman, T. Burr, D. Forslund, E. Joyce, R. Picard, and E. Umland. Modeling emergency department visit patterns for infectious disease complaints: results and application to disease surveillance. *BMC Medical Informatics and Decision Making*, 5 (1):4, 2005.

[99] D. Böhning. Zero-inflated Poisson models and C.A. MAN: A tutorial collection of evidence. *Biometrical Journal*, 40(7):833–843, 1998.

[100] D. Lambert. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14, 1992.

[101] G.J. McLachlan and D. Peel. *Finite mixture models*. Wiley-Interscience, 2000.

[102] S. Gurmu and P.K. Trivedi. Excess zeros in count models for recreational trips. *Journal of Business & Economic Statistics*, 14(4):469–477, 1996.

[103] E. Dietz and D. Böhning. On estimation of the Poisson parameter in zero-modified Poisson models. *Computational Statistics & Data Analysis*, 34(4):441–459, 2000.

[104] D.K. Agarwal, A.E. Gelfand, and S. Citron-Pousty. Zero-inflated models with application to spatial count data. *Environmental and Ecological Statistics*, 9(4):341–355, 2002.

[105] A.H. Lee, K. Wang, J.A. Scott, K.K.W. Yau, and G.J. McLachlan. Multi-level zero-inflated Poisson regression modelling of correlated count data with excess zeros. *Statistical methods in medical research*, 15(1):47, 2006.

[106] S.P. Miaou. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention*, 26(4):471–482, 1994.

[107] B. Crepon and E. Duguet. Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity. *Journal of Econometrics*, 79(2):355–378, 1997.

[108] D.C. Heilbron. Zero-altered and other regression models for count data with added zeros. *Biometrical Journal*, 36(5):531–547, 1994.

[109] V.T. Farewell and D.A. Sprott. The use of a mixture model in the analysis of count data. *Biometrics*, 44(4):1191–1194, 1988.

[110] D.Y.T. Fong and P. Yip. An EM algorithm for a mixture model of count data. *Statistics & Probability Letters*, 17(1):53–60, 1993.

[111] P.L. Gupta, R.C. Gupta, and R.C. Tripathi. Analysis of zero-adjusted count data. *Computational Statistics & Data Analysis*, 23(2):207–218, 1996.

[112] Y.B. Cheung. Zero-inflated models for regression analysis of count data: a study of growth and development. *Statistics in medicine*, 21(10):1461–1469, 2002.

[113] D. Böhning, E. Dietz, P. Schlattmann, L. Mendonça, and U. Kirchner. The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 162(2):195–209, 1999.

[114] J.D. Lewsey and W.M. Thomson. The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status. *Community Dentistry and Oral Epidemiology*, 32(3):183–189, 2004.

[115] K.K.W. Yau, K. Wang, and A.H. Lee. Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal*, 45 (4):437–452, 2003.

[116] S.M. Mwalili, E. Lesaffre, and D. Declerck. The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. *Statistical Methods in Medical Research*, 17(2):123, 2008.

[117] M. Minami, C.E. Lennert-Cody, W. Gao, and M. Román-Verdesoto. Modeling shark bycatch: the zero-inflated negative binomial regression model with smoothing. *Fisheries Research*, 84(2):210–221, 2007.

[118] J. Van den Broek. A score test for zero inflation in a Poisson distribution. *Biometrics*, 51(2):738–743, 1995.

[119] M. Ridout, J. Hinde, and C.G.B. Demétrio. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics*, 57(1):219–223, 2001.

[120] A.H. Lee, L. Xiang, and W.K. Fung. Sensitivity of score tests for zero-inflation in count data. *Statistics in medicine*, 23(17):2757–2769, 2004.

[121] P.L. Gupta, R.C. Gupta, and R.C. Tripathi. Score test for zero inflated generalized Poisson regression model. *Communications in Statistics-Theory and Methods*, 33(1): 47–64, 2004.

[122] M. Xie, B. He, and T.N. Goh. Zero-inflated Poisson model in statistical process control. *Computational Statistics & Data Analysis*, 38(2):191–201, 2001.

[123] M. Xie and T.N. Goh. SPC of a near zero-defect process subject to random shocks. *Quality and Reliability Engineering International*, 9(2):89–93, 1993.

[124] S. He, W. Huang, and W.H. Woodall. CUSUM charts for monitoring zero-inflated Poisson processes, to appear in Quality and Reliability Engineering International.

[125] M. Ridout, C.G.B. Demétrio, and J. Hinde. Models for count data with many zeros. In *Proceedings of the XIXth International Biometric Conference*, volume 19, pages 179–192, 1998.