Risk-Adjusted Exponentially Weighted Moving Average for Poisson Data and Application in Healthcare

Risk-Adjusted Exponentially Weighted Moving Average for Poisson Data and Application in Healthcare

By

Hui Wang

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2008) McMaster University

(Statistics) Hamilton, Ontario

TITLE: Risk-Adjusted Exponentially Weighted Moving Average for Poisson Data and Application in Healthcare

AUTHOR: Hui Wang, B.Sc. (McMaster University)

SUPERVISOR: Dr. Román Viveros-Aguilera

NUMBER OF PAGES: ix, 78

Acknowledgements

To finish this interesting project, I have received the whole-hearted support from my supervisor, Dr. Román Viveros-Aguilera. I am grateful for his professional advice, experience sharing, and encouragement which helped me in all the time of research and writing of this thesis.

My heartfelt appreciation also goes to the other members of the Examiner Committee. Dr. Angelo Canty and Dr. Rong Zhu, for their valuable suggestions on the content of my thesis. Especially, I would like to give my special thanks to Dr. Angelo Canty, who looked closely at the final version of the thesis for comments, questions, as well as English style and grammar, and offering suggestions for improvement.

I have furthermore to thank Dr. Peter Macdonald and Dr. Shui Feng for their teaching and instruction during the years I spend at McMaster University.

Finally, I would also like to dedicate this thesis to my wife, Lin Zhang and my sons, Ivan and Jonathan, as well as all my families whose patient love enabled me to complete this work.

Contents

List of Tables

List of Figures

Abstract

The Risk-Adjusted Exponentially Weighted Moving Average (RA-EWMA) for Poisson data is developed in detail. The method is useful to monitor healthcare or to other counts that are generated dynamically over time. The approach used is motivated by and follows closely the approach used by Grigg and Spiegelhalter (2007) for dynamic binary outcomes. In simple terms, a Bayesian approach is applied that uses conjugate priors (gamma in the case of Poisson data) utilized iteratively to provide the method estimates as the posterior expected means. The main application is counts with covariates. The thesis provides the necessary formulas to update the method's estimates. Numerical calculations are presented to illustrate the use of the methods and to compare it to the Standard Exponentially Weighted Moving Average (Standard EWMA), which is a standard monitoring method used in industrial applications. The numerical evidence provided in the thesis suggests that the RA-EWMA method is more sensitive than the Standard EWMA method to the presence of the underlying covariates. This was shown clearly on real data, specifically in the UK's death counts from lung diseases.

Chapter 1

Introduction and Thesis Objectives

1.1 Quality Control in Healthcare

The concepts and methods of total quality management (TQM) and continuous quality improvement (CQI) appeared right after World War II for improving the production quality of goods and services. They were not implemented in the healthcare area until recent years. There is now a growing demand in healthcare for the development of statistical process control (SPC) tools to measure and improve healthcare processes and outcomes (Carey and Lloyd, 2001).

Statistical process control techniques can be applied to different types of data such as clinical outcomes, risk management, and patient satisfaction. There is, however, a sharp distinguishing element between industrial and healthcare applications. For the most part, industrial settings involved production of items manufactured under controlled processes, yielding largely homogeneous product. Typical healthcare applications, on the other hand, while many aspects of the processes are under careful

supervision, the end receivers are patients presenting great diversity in their personal profiles. This diversity, which are collectively called *risk factors* in the literature, can have a substantive effect on the process outcome. The delay in the development of quality control methods in healthcare is likely in part due to the differences between the application areas. Taking the risk factors into account is a necessary condition for the success of any statistical method in healthcare.

1.2 Standard Monitoring

Patient safety is enhanced by the use of healthcare processes, working practices and systematic activities that prevent or reduce the risk of harm to patients. Patients achieve healthcare benefits that meet their individual needs through healthcare decisions and services, based on what assessed research evidence has shown provides effective clinical outcomes. Developing a proper flowchart to monitor the patient outcomes or the treatment procedure is important for patients to receive services as promptly as possible. The proper monitoring of a process can help healthcare providers choose the right services and treatments and avoid unnecessary delays at any stage of service delivery or the care pathway.

Statistics is a collection of techniques useful for making decisions about a process based on the analysis of the information collected. Statistical methods play a vital role in the quality control processes, they provide information used to control and improve the process. In statistical process control and monitoring, many of the techniques, such as the Shewhart control chart have been used for over 50 years. However, due to the increasing emphasis on variability reduction and process improvement, many

new statistical monitoring and control techniques have been developed. Examples include the cumulative sum (CUSUM) and the exponentially weighted moving average (EWMA) control charts. In general, a major disadvantage of the Shewhart control chart is that it only uses the information about the last data and ignores the entire previous history. This disadvantage makes the Shewhart control chart relatively insensitive to small shifts in the process. The CUSUM and EWMA control charts overcome this disadvantage and react quickly to even small changes. See Montgomery (2001) for details.

The CUSUM is a type of control chart based on the total deviations of successive samples from the target value. Each point plotted on the chart represents the sum of the deviations at the previous point, and all deviations since. It has been shown to be efficient in detecting small shifts in the mean of a process. If we resort to the traditional signal of an out-of-control process when one or more points fall beyond the control limit, then the Shewhart control chart might fail to detect the shift whereas the CUSUM control chart will detect it.

The exponentially weighted moving average is also a good method when we are interested in detecting small shifts. The performance of the EWMA is approximately the same as that of the CUSUM and it is easier to set up. The EWMA is usually used with individual observations.

The EWMA control chart was introduced by (Roberts 1959), and detailed discussions can be found in Crowder (1987) and in Lucas and Saccucci (1990). An exponentially weighted moving average applies weighting factors to the data points. The weighting for each data point decreases exponentially, giving much more importance to recent observations while still not discarding older observations entirely. It

is a well known method for time series forecasting and smoothing. The EWMA can be expressed as the following form by Harvey (1991)

$$Current\ estimate = Previous\ estimate + discounted\ error.$$

Hence is one type of Kalman filter. By enclosing the state-space models, this filtering method can be extended to deal with the estimation with covariates by Harvey (1991) and West and Harrison (1997).

As with other control charts, both CUSUM and EWMA charts are used to monitor processes over time. The charts are time based so that they show a history of the process.

The degree of weighing decrease is expressed as a constant smoothing factor κ, a number between 0 and 1. κ may be expressed as a percentage, so a smoothing factor of 5% is equivalent to $\kappa = 0.05$. The EWMA statistic provides a smoothed estimate of the current level of the process. Because the EWMA chart uses information from all samples, the prediction for next outcome or outcomes will depend on previous observations, hence the response will be fast when the the shift in the mean of a process occurs.

Recently, increased interest has been placed in monitoring heterogeneous time series in medical contexts. The aim is to monitor and control the outcome of a medical procedure or process. The risk factors (i.e. patient covariates) have potentially an effect on the procedure outcomes. In general, the traditional EWMA method as described above treats all patients as having equal risk factors. This makes the methods ineffective in some cases and misleading in others when monitoring healthcare outcomes. For instance, a sudden increase in the number of failures in the outcomes

may be due to the treatment of several high risk patients and not to a change in the application of the healthcare service. The result is unnecessary false alarms in the monitoring. Likewise, treating several low-risk cases yielding an unalarming small number of failures may result in an undetected deterioration of the service when using the traditional EWMA. Grigg and Spiegelhalter (2007) have addressed these problems and have developed several promising approaches to incorporate risk factors in process monitoring and control of healthcare processes. Specifically, they provide a thorough analysis of processes with binary (i.e. success/failure) outcomes.

In the medical context, count data often come up in a systematic and regular way. For example, each month the number of patients who die due to lung diseases in a particular area. These outcomes clearly have seasonal as well as other trends. If we want to discover the secrets behind process changes over time, we need to consider temperature, pollution as well as other factors which will affect the patients.

1.3 Thesis Objectives and Overview

In this thesis, we focus on the exponentially weighted moving average control chart with risk-adjustment which emphasis on the case of Poisson data. Specifically we aim to develop:

- the theoretical aspects for the Poisson case;
- its numerical implementation; and
- a comparison with competing charts through simulations.

In Chapter 2, we review the standard and risk-adjusted EWMA methods for binary data as developed by Grigg and Spiegelhalter (2007), and reproduce some of the simulation studies and plots to illustrate the basic ideas of these two methods. In Chapter 3, we make the theoretical development of the EWMA and RA-EWMA methods for Poisson data. In Chapter 4, we first make a simulation-based comparison between EWMA and RA-EWMA for Poisson data, then apply the methods to the UK lung diseases death data and compare the results. Finally, we make concluding remarks and discuss further possible extensions in Chapter 5.

Chapter 2

Standard and Risk-Adjusted EWMA for Binary Data

In some medical applications, there is an interest in monitoring health outcomes over time while taking the severity of individual patient's condition into account. In general, the severity and other patient covariates are called risk factors. In their research, Grigg and Spiegelhalter (2007) provided a direct estimate of the current chance of an adverse event given a patient's covariate information. As we mentioned in Chapter 1, the order of patient arrival might mask the true change in the hidden risk for a patient. Adjusting for the different effects exerted by the risk factors for the patient will render more accurate information for the healthcare processes.

In this chapter we review the risk-adjusted exponentially weighted moving average method for binary data developed by Grigg and Spiegelhalter (2007). Section 2.1 introduces the basic idea of the standard exponentially weighted moving average smoothing method for binary data. Section 2.2 introduces the EWMA for the

Bayesian state-space model. Section 2.3 presents a comparison between the dynamic generalized linear model and the mean steady model. A simulation study is presented in Section 2.4 focusing on the performance of the standard EWMA and the risk-adjusted EWMA methods. Concluding remarks regarding the performance of the methods for binary outcomes are summarized at the end of the simulation study.

2.1 The Exponentially Weighted Moving Average

The standard exponentially weighted moving average is a well-known estimation and prediction tool in time series analysis. It has also being applied extensively in statistical process control and monitoring particularly to industrial processes (Montgomery 2001). The basic form of the method is as follows.

Let $y_1, ..., y_t, ...$ be a sequence of random variables at successive times. Assume that $\mathbb{E}[y_t|\mu_t] = \mu_t$ for $t = 1, 2, ...$, where t is the index of time and μ_t's are assumed to be from some dynamic process. The standard exponentially weighted moving average based on the observations $y_1, ..., y_t$ is the statistic $\hat{\mu_t}^E$ defined recursively by

$$\begin{aligned}
\hat{\mu}_t^E &= \kappa\hat{\mu}_{t-1}^E + (1-\kappa)y_t \quad (t = 1, 2, ..., \ 0 \leq \kappa \leq 1) \qquad (2.1.1)\\
&= \kappa(\kappa\hat{\mu}_{t-2}^E + (1-\kappa)y_{t-1}) + (1-\kappa)y_t \\
&= \kappa^2\hat{\mu}_{t-2}^E + (1-\kappa)(\kappa y_{t-1} + y_t) \\
&= \cdots \\
&= \kappa^t\hat{\mu}_0^E + (1-\kappa)\sum_{i=1}^{t}\kappa^{t-i}y_i, \qquad (2.1.2)
\end{aligned}$$

where $\hat{\mu}_0^E$ is the estimate for μ_0. In general, the $\hat{\mu}_0^E$ is derived from the training data or early data by the given model.

The most frequent use of $\hat{\mu}_t^E$ is as an estimator of the process level at time t, namely μ_t. The parameter κ determines the speed of decay. In the case of $\kappa = 0$, the current estimate of the mean depends only on current data, and if $\kappa = 1$, the current estimate of the mean is totally dependent on the prior estimate of the mean. In the standard analysis without risk factor adjustment, $\hat{\mu}_t^E$ is a forecast for μ_{t+1}.

Based on the standard EWMA, the risk-adjusted exponentially weighted moving average (RA-EWMA) provides a smoothed estimate for the expected outcome for an observation taking into account its covariate values. For example, if we monitor the 30-day mortality of patients after cardiac surgery by a particular surgeon, the patient's gender, age, diabetes status, and other preoperative factors can be considered as covariates for the patient (Steiner *et al.*, 2000). The derivation given in this section, the adjustment made for current differential risk is based on an approximation to the likelihood, which is assumed to be of exponential family form.

Assuming a distribution from the exponential family, the probability mass function or probability density function of y_t can be written as

$$f(y_t|\eta_t^+, \phi) = u(y_t, \phi) \exp\{[y_t\eta_t^+ - \nu(\eta_t^+)]/\phi\}, \quad t = 1, 2, ..., \tag{2.1.3}$$

where η_t^+ is the natural parameter, we have $g(\mu_t^+) = \eta_t^+$ and $\mathbb{E}[y_t|\mu_t^+] = \mu_t^+$, where $g(\cdot)$ is the canonical link function. " + " indicates that a parameter includes the effect of risk factors. ϕ is assumed to be a known scale parameter, namely dispersion parameter.

From the Generalized Linear Model (GLM), we have $\mathbb{E}[y_t|\mu_t^+] = \nu'(\eta_t^+)$ and $\mathbb{V}[y_t|\mu_t^+] = \nu''(\eta_t^+)\phi$ by (McCullagh and Nelder 1983), where $\nu'(\cdot)$ is the first derivative of $\nu(\cdot)$ and $\nu''(\cdot)$ is the second derivative of $\nu(\cdot)$ with respect to the unknown

parameter.

For example, assume

$$y_t|\mu_t \sim Bernoulli(\mu_t),\ t = 1, 2, ...$$

The probability mass function can be given by

$$\begin{aligned} f(y_t|\mu_t) &= \mu_t^{y_t}(1-\mu_t)^{1-y_t} \\ &= \exp\{y_t \log(\mu_t) + (1-y_t)\log(1-\mu_t)\} \\ &= \exp\{y_t \log(\frac{\mu_t}{1-\mu_t}) + \log(1-\mu_t)\} \\ &= u(y_t, \phi)\exp\{\phi^{-1}[y_t\eta_t - \nu(\eta_t)]\}, \end{aligned}$$

where $u(y_t, \phi) = 1$ and $\phi = 1$. In above equation, we have

$$\eta_t = \log(\frac{\mu_t}{1-\mu_t})$$

and

$$\nu(\eta_t) = -\log(1-\mu_t).$$

Now we assume a structure for η_t^+, a natural parameter includes the effects of covariates

$$\eta_t^+ = \eta_t + \delta_t,\ \ t = 1, 2, ..., \tag{2.1.4}$$

where δ_t denote the risk adjustment level at time t associated with observation y_t, $t = 1, 2, ...,$. Assume that δ_t in equation (2.1.4) has the from $\delta_t = \boldsymbol{\beta}^T \boldsymbol{x_t}$, where $\boldsymbol{x_t}$ be a vector of observed and centered covariates, thus $\boldsymbol{x_t} = \mathbf{0}$ is the baseline. In addition, we consider the vector of coefficients $\boldsymbol{\beta}$ to be known usually from fitting the model to data gathered when the healthcare process was operating in control. Also,

the baseline expectation is taken to be $\mu_t = g^{-1}(\eta_t)$. Hence equation (2.1.4) can be rewritten as

$$g(\mu_t^+) = g(\mu_t) + \delta_t, \quad t = 1, 2, ... \tag{2.1.5}$$

Using the standard GLM notation, the log-likelihood function arising from equation (2.1.3) can be written as

$$\begin{aligned} L_1(\mu_t) &= \{y_t\eta^+ - \nu(\eta^+)\}/\phi \\ &= \{y_t[g(\mu_t) + \delta_t] - \nu[g(\mu_t) + \delta_t]\}/\phi. \end{aligned} \tag{2.1.6}$$

Even though δ_t is known, equation (2.1.6) is still complicated for inferences about μ_t. If we use approximate analysis instead of the exact analysis, the form of the expectation of μ_t might be easier. Let y_t^p be a risk-adjusted pseudo-baseline observation, assume it stratifies $\mathbb{E}[y_t^p|\mu_t] = \mu_t$, then the log-likelihood function of equation (2.1.3) would be

$$L_2(\mu_t) = \{y_t^p[g(\mu_t)] - \nu[g(\mu_t)]\}/\phi. \tag{2.1.7}$$

What we want is to obtain a statistic y_t^p whose contribution to the likelihood of μ_t is equal to that made by the original data y_t. Then the inference of μ_t based on equation (2.1.7) would be the same as that based on equation (2.1.6) for all t. Thus the likelihood contributions made by y_t^p and y_t would be exactly equal for any true value of μ_t if the score functions of (2.1.7) and (2.1.6) were identical at all μ_t. The score function can be written as

$$L_1' = g'(\mu_t)(y_t - \mu_t^+)/\phi \tag{2.1.8}$$

and

$$L_2' = g'(\mu_t)(y_t^p - \mu_t)/\phi. \tag{2.1.9}$$

The identity occurs if

$$L_1' = L_2' \tag{2.1.10}$$

hence we have

$$y_t^p = y_t - (\mu_t^+ - \mu_t). \tag{2.1.11}$$

This equation indicates that the pseudo-observation needs to be the difference between original observation and its differential expectation $\mu_t^+ - \mu_t$. We can use an estimated value to replace the unknown parameters. In this case, (2.1.11) can be replaced by

$$\tilde{y}_t = y_t - (\hat{\mu}_t^+ - \hat{\mu}_t) \tag{2.1.12}$$

where

$$\hat{\mu}_t^+ = g^{-1}(g(\hat{\mu}_t) + \delta_t). \tag{2.1.13}$$

In the same way as in the standard EWMA, the RA-EWMA at time $t-1$ can provide a forecast for the baseline expectation μ_t at time t.

If we wish to estimate the baseline mean parameter μ_t of the exponential family data y_t, given risk-adjustment level δ_t, we can define a RA-EWMA as

$$\hat{\mu}_t^R = \kappa\hat{\mu}_{t-1}^R + (1-\kappa)\tilde{y}_t \quad (t = 1, 2, ..., \ 0 \leq \kappa \leq 1) \tag{2.1.14}$$

$$= \kappa^t\hat{\mu}_0^R + (1-\kappa)\sum_{i=1}^{t}\kappa^{t-i}\tilde{y}_i, \tag{2.1.15}$$

where $\hat{\mu}_0^R$ is an estimate of μ_0. In general, the $\hat{\mu}_0^R$ is derived from the training data or early data by the given model.

Noticing from equation (2.1.2) that $\hat{\mu}_t^E$ is a linear combination of $\hat{\mu}_0^E$, y_1, y_2, ..., y_t and these variables are independent it follows that the variance for the standard EWMA is

$$\mathbb{V}[\hat{\mu}_t^E] = (1-\kappa)^2 \sum_{i=1}^{t} \kappa^{2(t-i)} \mathbb{V}[y_i]. \tag{2.1.16}$$

Note that $\mathbb{V}[\hat{\mu}_0^E]$ is taken to be 0. If the dispersion parameter ϕ of the y_t is assumed fixed and known, then using the basic finite geometric sum the above equation simplifies to

$$\mathbb{V}[\hat{\mu}_t^E] = \frac{1-\kappa}{1+\kappa}(1-\kappa^{2t})\phi, \tag{2.1.17}$$

which has limiting value $\sigma^2 = \phi(1-\kappa)/(1+\kappa)$ as $t \to \infty$.

2.1.1 Example: Mortality Rate After Cardiac Surgery

In this section we present an example aimed at illustrating the implementation of the RA-EWMA and also to compare it with the standard EWMA. The data collected are the 30-day mortality after cardiac surgery in a UK Cardiac Surgery Center by one of the surgeons over the years 1992 and 1998. Specifically,

$$y_t = \begin{cases} 1 \text{ if the } t\text{th patient died within 30 days after surgery;} \\ 0 \text{ otherwise.} \end{cases}$$

Where t is patient number in the order the patients were treated and y_t is the associated outcome. Thus the probability mass function here is simply the Bernoulli mass function

$$f(y_t; p_t) = p_t^{y_t}(1-p_t)^{1-y_t} \quad (t = 1, 2, ..., \ y_t = 0, 1).$$

In this example, all the required data and preliminary analysis results were given by Steiner *et al.* (2000). Our main objective is to show in detail how the the RA-EWMA is implemented.

A logistic regression model has been fitted to develop the risk equation. To fit this model, 2,218 operations during year 1992 and 1993 have been choose as training data. The resulting equation is given by Grigg and Spiegelhalter (2007)

$$\eta_t^+ \approx \eta_t + .077x_t,$$

where x_t is the centered Parsonnet score for acquired adult heart surgery, standardized to have mean zero in the training data. Originally introduced by Parsonnet (1989), the score is a preoperative predictor of mortality in cardiac surgery based on a variety of risk factors that include gender, age, diabetes status, number of catastrophic states, and others. The Parsonnet score has become very popular in heart studies in recent years. Hence the risk adjustment level is known, which is $\delta_t = \beta x_t = .077x_t$ for the t^{th} patient. For simplicity, we choose 5 consecutive surgical outcomes for the year 1994-1998 under a particular surgeon. We denote the outcomes by vector $\mathbf{y}$ and $\mathbf{y} = (1, 0, 1, 0, 1)^T$. The covariate vector is given by $\mathbf{x} = (44, -6, 22, 15, 42)^T$. Hence the risk factor vector for the given patients can be calculated as $\boldsymbol{\delta} = \beta\mathbf{x} = (3.4, -.5, 1.7, 1.2, 3.2)^T$.

For the initial setup, the training data have 2,218 observations, 142 patients died within 30 days after surgery, so this gives $\mu_0^E = 142/2218 = .064$. If we using the decay parameter $\kappa = 0.9$, then equation (2.1.2) gives the EWMA as $\hat{\mu}^E = (.158, .142, .228, .205, .285)^T$. In the estimation of the standard EWMA, we simply set the risk adjustment level to be zero for each of the observation.

Next, we are going to calculate the RA-EWMA. Starting with an estimate for the natural parameter value of η_0 of $\eta_0 = -3.0$, which is equal to the intercept value of the generalized linear model fitted to the original data. We have $\hat{\mu}_0^R = e^{\hat{\eta}_0}/(1+e^{\hat{\eta}_0}) \approx$

0.048. To calculate $\hat{\mu}_1^+$, we have

$$\hat{\mu}_1^+ = g^{-1}(g(\hat{\mu}_0^R) + \delta_1) = .596.$$

Plug in

$$\tilde{y}_1 = y_1 - (\hat{\mu}_1^+ - \hat{\mu}_1),$$

we get $\tilde{y}_1 = 0.451$. Next, by equation (2.1.14),

$$\hat{\mu}_1^R = \kappa\hat{\mu}_0^R + (1-\kappa)\tilde{y}_1 = .088.$$

By above recursive method, we finally get the EWMA and RA-EWMA values for $\mathbf{y}$ in Table 2.1.

The EWMA estimates the current mortality probability for the t^{th} patient and the RA-EWMA provides an estimate of mortality probability for the t^{th} patient via $\hat{\mu}_t^+ = g^{-1}(g(\hat{\mu}_{t-1}^R) + \delta_t)$, where the risk adjustment level is given by $\delta_t = .077x_t$. It is important to clarify that for a non-linear link function, for example, the log link function for Poisson data and the logistic link function for the Bernoulli data, the EWMA and the risk-adjusted EWMA are estimating different quantities. Precisely, the EWMA is estimating the risk parameter under the assumption of homogeneity of patients, but the RA-EWMA is estimating the risk parameter for a standard patient with zero as the centered Parsonnet score (predictive score for acquired adult heart surgery). This is the reason why these two methods use different starting values. Only when risk adjustment factors is factored into the RA-EWMA to provide a patient-specific risk assessment can we reasonably make a comparison with the standard EWMA.

Table 2.1: *The estimated values by EWMA and RA-EWMA methods for cardiac surgery outcomes* y_t, $t = 1, ..., 5$, *under a particular surgeon in the UK.*

t	$\hat{\mu}_t^E$	$\hat{\mu}_t^R$	y_t	x_t	$\hat{\mu}_t^+$	$\tilde{y}_t$
0	.064	.048				
1	.158	.088	1	44	.596	.451
2	.142	.082	0	-6	.058	.030
3	.228	.149	1	22	.328	.754
4	.205	.113	0	15	.359	-.210
5	.285	.137	1	42	.764	.349

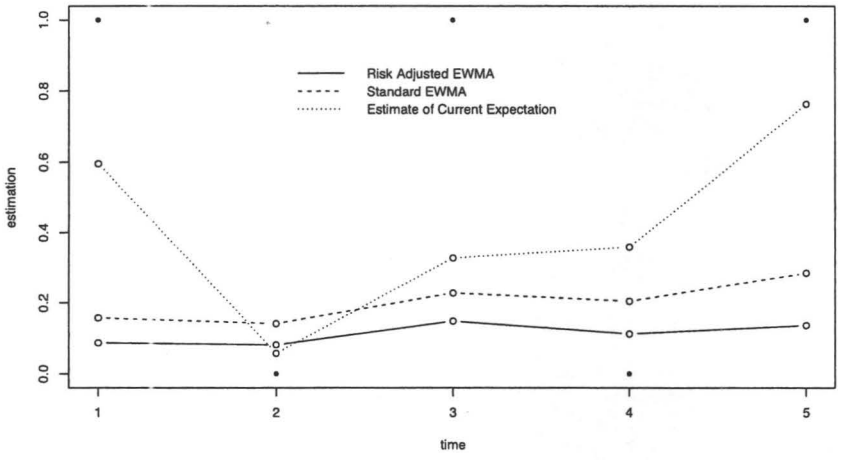

Figure 2.1: *Graphical display of the estimated values by EWMA and RA-EWMA methods for the Cardiac Surgery outcomes as given in Table 2.1. The black dots are the outcome* y_t *values.*

Table 2.1 gives the estimated values of standard EWMA ($\hat{\mu}_t^E$), RA-EWMA ($\hat{\mu}_t^R$), and the estimate of the current expectation $\hat{\mu}_t^+$. Comparing the values of $\hat{\mu}_t^+$ and $\hat{\mu}_t^E$ in Table 2.1, we see a sizable difference. This difference is because of the differential risks for the individual patients, some exhibiting Parsonnet scores far away from the baseline. For example, the first and the last patients have high risk factor levels. Note that the estimate of the current expected value, μ_t^+, is very high for these two cases but the RA-EWMA estimates mitigate their effect.

Figure 2.1 shows graphically the estimated values of standard EWMA $\hat{\mu}_t^E$, RA-EWMA $\hat{\mu}_t^R$, and the estimate of the current expectation $\hat{\mu}_t^+$. Note that, both in Figure 2.1 and Table 2.1, the standard EWMA estimates the current mortality probability for the tth patient, the RA-EWMA estimates the current mortality probability for a baseline patient with $x_t = 0$.

2.2 EWMA from Conjugate State-Space Models

Denote the observed data up to time t by $d_t = (y_1, y_2, ..., y_t)$. In our previous analyses, we took the viewpoint that μ_t was fixed and unknown, and that it could be consistently estimated in repeated experiments. In this section we take a Bayesian approach and work with the estimate $\hat{\mu}_t$ given by the posterior mean of μ_t.

There is a tradition of deriving EWMAs from state-space models, in which the data d_t up to time t provide a posterior distribution $p(\mu_t|d_t)$ for the mean parameter μ_t. A dynamic evolution process gives a forecast prior distribution $p(\mu_{t+1}|d_t)$. After we have the new observation y_{t+1}, a Bayesian conjugate update is given by $p(\mu_{t+1}|d_{t+1}) \propto p(y_{t+1}|\mu_{t+1})p(\mu_{t+1}|d_t)$. The evolution process can be defined on the natural or mean

parameter scale. The following formula is defined as natural scale by West and Harrison (1997),

$$p(\eta_t|d_t) \propto \exp[r_t\eta_t - s_t\nu(\eta_t)], \tag{2.2.1}$$

denoted simply as $\eta_t|d_t \sim CP(r_t, s_t)$ for the conjugate prior on the natural scale, where s_t can be interpreted as precision parameter, the reciprocal of the dispersion parameter, and r_t/s_t can be interpreted as location parameters. In certain situations, the mean scale $\mu_t|d_t \sim CP(r_t, s_t)$ is more convenient. It has mean $\mathbb{E}[\mu_t|d_t] = r_t/s_t$ and $\mathbb{V}[\mu_t|d_t] = \mathbb{E}[\nu''(\eta_t)|d_t]/s_t$. The precision s_t is denoted as $\mathbb{P}[\mu_t|d_t]$. When the new observation y_{t+1} comes in, the posterior distribution can be expressed as

$$\mu_t|d_t, y_{t+1} \sim CP(r_t + \phi^{-1}y_{t+1}, s_t + \phi^{-1}). \tag{2.2.2}$$

The mean steady model (MSM) leads the RA-EWMA as an estimator for the current mean μ_t. We assume a mean steady evolution step, the forecasted mean is equal to the posterior mean but the precision is decreased by a factor κ. A posterior distribution $\mu_t|d_t \sim CP(r_t, s_t)$ leads to a forecast prior $\mu_{t+1}|d_t \sim CP(\kappa r_t, \kappa s_t)$. By the above assumptions, the update step is exact, the risk-adjustment step can be made only approximately, and the evolution step is justifiable only heuristically.

We then have the following properties

$$\begin{aligned}\mathbb{E}[\mu_{t+1}|d_t] &= \mathbb{E}[\mu_t|d_t] = \frac{r_t}{s_t},\\ \mathbb{P}[\mu_{t+1}|d_t] &= \kappa\mathbb{P}[\mu_t|d_t] = \kappa s_t.\end{aligned} \tag{2.2.3}$$

For the case of Bernoulli data, an approximate risk adjustment can be made after the new observation y_{t+1} comes by using the pseudo-observation $\tilde{y}_{t+1}$ in the conjugate

update of μ_{t+1}, so that we have

$$\begin{aligned}\mu_{t+1}|d_{t+1} &\sim CP(\kappa r_t + \phi^{-1}\tilde{y}_{t+1}, \kappa s_t + \phi^{-1}) \\ &\sim CP(r_{t+1}, s_{t+1}).\end{aligned} \tag{2.2.4}$$

In the limiting form as $t \to \infty$, the posterior expectation and precision are given by

$$\begin{aligned}\mathbb{E}[\mu_{t+1}|d_{t+1}] &= \kappa\mathbb{E}[\mu_t|d_t] + (1-\kappa)\tilde{y}_{t+1}, \\ \mathbb{P}[\mu_{t+1}|d_t] &= s = \phi^{-1}(1-\kappa)^{-1};\end{aligned} \tag{2.2.5}$$

that is, the posterior expectation tends to an EWMA on the $\tilde{y}_t$'s.

2.3 Comparison Between DGLM and MSM

The Dynamic Generalized Linear Model (DGLM) is often considered the natural method for nonnormal data with covariates, but it is overcomplicated for many contexts. In this section, we give a comparison between the DGLM and the MSM for Bernoulli data.

The DGLM is a Bayesian forecasting method for time series data when the risk factors have been described by a generalized linear model structure McCullagh and Nelder (1983). The parameters of the DGLM are allowed to change over time. In general, the evolution step and risk adjustment step are exact, but the parameter update step is approximate.

2.3.1 Dynamic Generalized Linear Model

The following steps outline the procedure for Bernoulli data.

1. **Conjugate baseline mean**. Assume a conjugate distribution for μ_t,

$$\mu_t|d_t \sim CP(r_t, s_t), \quad t = 1, 2, ..., \tag{2.3.1}$$

which for Bernoulli data will be $\beta(r_t, s_t)$.

2. **Natural baseline**. Translate the above conjugate distribution to the natural scale for $\eta_t|d_t$ with $m_t = \mathbb{E}[\eta_t|d_t]$ and $C_t = \mathbb{V}[\eta_t|d_t]$ by the transformation equation $\eta_t = g(\mu_t)$. For Bernoulli distribution, the link function is $\eta_t = \log(\mu_t/(1-\mu_t))$, the results are given by

$$\begin{aligned} m_t &= \gamma(r_t) - \gamma(s_t - r_t) \approx \log \frac{r_t}{(s_t - r_t)}, \\ C_t &= \gamma'(r_t) - \gamma'(s_t - r_t) \approx \frac{1}{r_t} + \frac{1}{(s_t - r_t)}, \end{aligned} \tag{2.3.2}$$

where $\gamma(\cdot)$ and $\gamma'(\cdot)$ represent the digamma and trigamma function respectively. Now we can assume that the distribution $\eta_t|d_t$ has the required mean and variance, but the precise form is left unspecified.

3. **Evolution**. The process evolves by adding a term with a distribution $[0, W]$, with mean zero and variance W, but the form unspecified. This evolution will give a forecast prior distribution for $\eta_{t+1}|d_t$, with

$$\eta_{t+1}|d_t \sim [m_t, C_t + W] \tag{2.3.3}$$

4. **Risk-adjustment**. Given the risk term δ_{t+1} at time $t+1$, we then have

$$\begin{aligned} \eta^+_{t+1}|d_t, \delta_{t+1} &\sim [m_t + \delta_{t+1}, C_t + W] \\ &\sim [e_t, q_t] \end{aligned} \tag{2.3.4}$$

5. **Transform to mean scale distribution**. A conjugate prior distribution can be obtained by the inverse transformation of the second stage. The distribution is

given by

$$\mu_{t+1}^{+}|d_t, \delta_{t+1} \sim CP(r_t^*, s_t^*), \tag{2.3.5}$$

where

$$r_t^* \approx \frac{1 + e^{e_t}}{q_t}$$

and

$$s_t^* \approx \frac{2 + e^{e_t} + e^{-e_t}}{q_t}.$$

6. **Conjugate update on mean scale**. After we have observation y_{t+1}, we have the posterior distribution

$$\mu_{t+1}^{+}|d_{t+1} \sim CP(r_t^* + \phi^{-1}y_{t+1}, s_t^* + \phi^{-1}), \tag{2.3.6}$$

7. **Transform to natural scale and de-risk-adjust**. If we let

$$r_t^{**} = r_t^* + \phi^{-1}y_{t+1}$$
$$s_t^{**} = s_t^* + \phi^{-1},$$

an approximate posterior on the natural scale can be obtained by using the transformation in Stage 2. The transformation is as

$$e_t^* \approx \log \frac{r_t^{**}}{(s_t^{**} - r_t^{**})},$$
$$q_t^* \approx \frac{1}{r_t^{**}} + \frac{1}{(s_t^{**} - r_t^{**})}.$$

Hence, we have the updated natural scale expression

$$\eta_{t+1}^{+}|d_{t+1} \sim [e_t^*, q_t^*], \tag{2.3.7}$$

then de-risk

$$\begin{aligned}\eta_{t+1}|d_{t+1} &\sim [e_t^* - \delta_{t+1}, q_t^*] \\ &\sim [m_{t+1}, C_{t+1}].\end{aligned} \tag{2.3.8}$$

8. **Translate to conjugate baseline.** Using the inverse transformation of Stage 5, we can obtain

$$\mu_{t+1}|d_{t+1} \sim CP(r_{t+1}, s_{t+1}), \quad t = 1, 2, ..., \tag{2.3.9}$$

where

$$r_{t+1} \approx \frac{1 + e^{m_{t+1}}}{C_{t+1}}$$

and

$$s_{t+1} \approx \frac{2 + e^{m_{t+1}} + e^{-m_{t+1}}}{C_{t+1}}.$$

The whole process, from step 1 to step 8, is a closed form in terms of the inputs r_t, s_t, and y_t at time t and the outputs r_{t+1} and s_{t+1}. After this whole process, it is ready for next cycle in terms of the inputs r_{t+1}, s_{t+1}, and y_{t+1} at time $t+1$.

2.3.2 Mean Steady Model

The following steps outline the Mean Steady Model (MSM) procedure to derive the RA-EWMA for the Bernoulli example.

1. **Conjugate baseline mean.**

$$\mu_t|d_t \sim CP(r_t, s), \quad t = 1, 2, ..., \tag{2.3.10}$$

where $s = \phi^{-1}(1 - \kappa)^{-1}$ by previous equation.

2. **Evolve as a MSM.**

$$\mu_{t+1}|d_t \sim CP(\kappa r_t, \kappa s), \quad t = 1, 2, ..., \tag{2.3.11}$$

3. **Update using approximate risk-adjustment.** Using the pseudo-observation $\tilde{y}_{t+1}$ from equation (2.1.12), hence we have

$$\begin{aligned} \mu_{t+1}|d_{t+1} &\sim CP(\kappa r_t + \phi^{-1}\tilde{y}_{t+1}, \kappa s + \phi^{-1}) \\ &\sim CP(r_{t+1}, s). \end{aligned} \tag{2.3.12}$$

From above equation, we can see that

$$\begin{aligned} \hat{\mu}_{t+1} &= (\kappa r_t + \phi^{-1}\tilde{y}_{t+1})/s \\ &= (\kappa r_t)/s + (\phi^{-1}\tilde{y}_{t+1})/s \\ &= \kappa(r_t/s) + (\phi^{-1}/s)\tilde{y}_{t+1} \\ &= \kappa\hat{\mu}_t + (1-\kappa)\tilde{y}_{t+1} \end{aligned}$$

since $\hat{\mu}_t = r_t/s$ and $1 - \kappa = 1/(\phi s)$, immediately giving the form of the RA-EWMA.

To compare the DGLM and MSM methods, we need a mapping between the evolution parameter W of DGLM and the smoothing parameter κ of MSM.

Under a DGLM, the posterior variance of the natural parameter can be given by the delta method as

$$C_t = \mathbb{V}[\eta_t|d_t] \approx \left.\frac{\mathbb{V}[\mu_t|d_t]}{(d\mu_t/d\eta_t)^2}\right|_{\eta_t=\hat{\eta}_t}, \tag{2.3.13}$$

where $\hat{\eta}_t$ is the MLE for η_t. We also can find the relation between the posterior variance of the mean parameter and posterior precision parameter s by West and Harrison (1997), the relation can be written as

$$\mathbb{V}[\mu_t|d_t] = \mathbb{E}[\nu''(\eta_t)|d_t]/s \approx \nu''(\hat{\eta}_t)/s, \tag{2.3.14}$$

and $d\mu_t/d\eta_t = \nu''(\eta_t)$. Thus from (2.3.13) and (2.3.14), we can approximate the posterior variance by $C_t \approx 1/(s\nu''(\hat{\eta}_t))$.

2.4 Simulation Study for Binary Data

In this section, we present a simulation study for Bernoulli data to compare the estimated values by the DGLM and MSM methods. The EWMA been derived by MSM method when set the risk-adjustment level $\delta_t = 0$ for all t and the RA-EWMA been derived when not all $\delta_t = 0$.

In the first example, the risk factors have been set to zeros. In other words, let $\delta_t = 0,\ t = 1, ..., T$ for all the observations. Let $\mu_0 = 0.2$, and $W = 0.03$. With these settings, $\hat{\nu}_t = \mu_0(1 - \mu_0) = 0.16$, Assuming $\phi = 1$ and solving $W \approx \phi(1 - \kappa)^2/(\kappa\hat{\nu}_t)$, we get approximately $\kappa = 0.93$ as a reasonable approximation.

For the Bernoulli data simulation, we setup the process as two stages. First, the trendless baseline natural parameter η_t, $t = 1, ..., 60$, were simulated as the training data by the distribution $\eta_t|\eta_{t-1} \sim N(\eta_{t-1}, W)$. For $t = 61, ..., 260$, $\eta_t|\eta_{t-1} \sim N(\rho\eta_{t-1}, W)$, with shift parameter $\rho = 0.99$ as the test data.

In the second stage, using the simulated η_t, we calculate μ_t by using the link function $g(\cdot)$, where $g(\cdot)$ for Binary data is the logistic function

$$\eta_t = \frac{log(\mu_t)}{1 - log(\mu_t)}.$$

The homogeneous data were generated by sampling from Bernoulli(μ_t), and the heterogeneous data were generated by sampling from Bernoulli(μ_t^+). Where $\mu_t^+ = g^{-1}[g(\mu_t) + \delta_t] = \exp(\eta_t)/[1 + \exp(\eta_t)]$. In this study, the risk-adjusted levels were

generated from Normal distribution with zero mean and variance 0.25. For the relative accuracy, each estimated series has been recorded for the above two methods, and then averaged over 500 simulations under each methods.

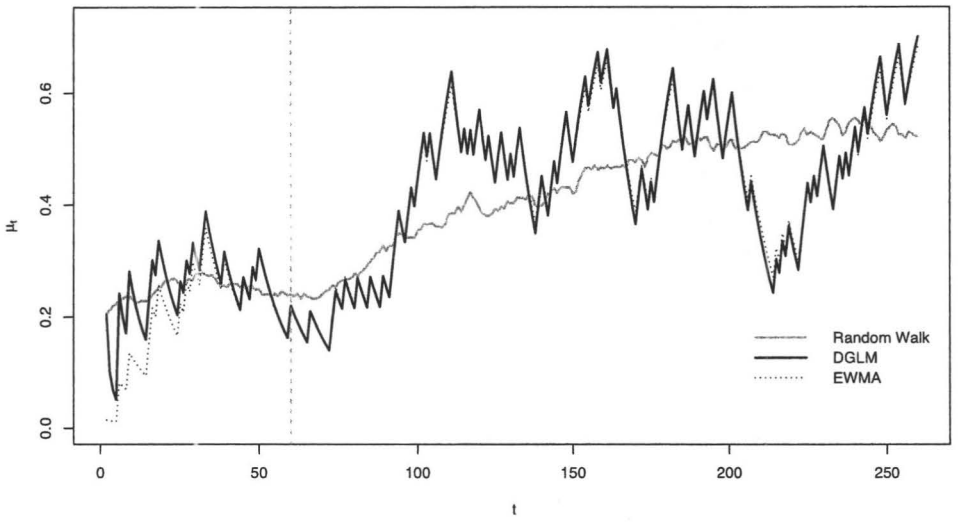

Figure 2.2: *Simulated Binary homogeneous data. The estimated series by EWMA and DGLM methods are displayed whereas the vertical line at* $t = 60$ *marks the time the random walk begins to shift.*

Figure 2.2 and 2.3 show plots of the one-step predictions for the DGLM and RA-EWMA with and without risk-adjustment. For the DGLM, equation (2.3.5) gives the one step predictive distribution and for RA-EWMA, equation (2.3.11) gives the one step predictive distribution. Only at the very beginning, the one step prediction shows difference between those two methods, all others are very similar.

In Figure 2.2, the model assumes homogeneity. The estimated values by both

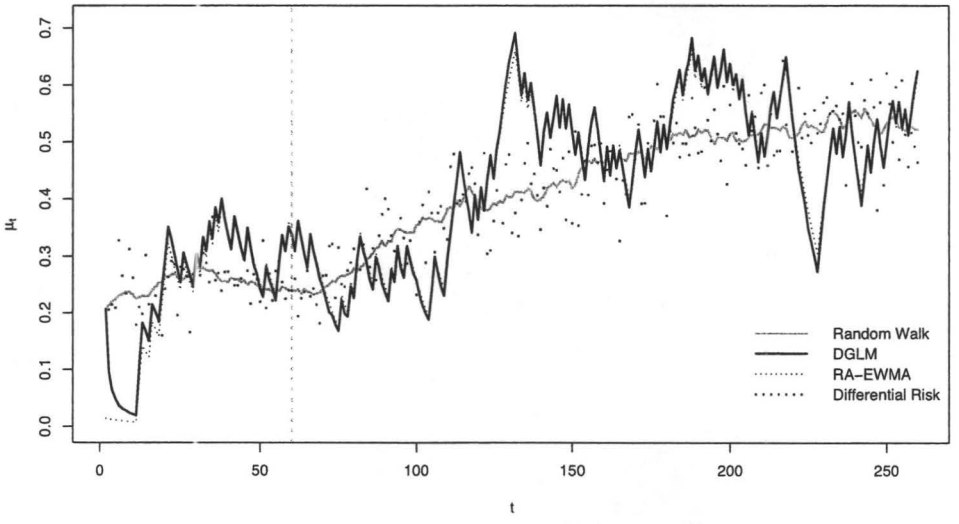

Figure 2.3: *Simulated Binary heterogeneous data. The estimated series by RA-EWMA and DGLM methods are displayed whereas the vertical line at* $t = 60$ *marks the time the random walk begins to shift. The points denote the expectation of each observation conditional on their risk-adjustment level* δ_t

DGLM and RA-EWMA methods are captured the trend of the random walk, which is an unknown dynamic process we want to estimate. At the training stage, t from 1 to 60, the estimated series by DGLM method appears captured the random walk faster comparing with the estimated series by RA-EWMA method. In the test stage, $t > 60$, these two methods perform almost identical except at the some peak or foot points along the estimated series. The vertical dish line at $t = 60$ separates the training data and test data.

In Figure 2.3, the model assumes heterogeneity. The dots denote the differential risk, which indicate the expectation of each observation conditional on their risk-adjustment level δ_t. Looking at the estimated results, they have same properties as those in Figure 2.2.

Traditionally, the DGLM has been considered as a standard method for dynamic monitoring under nonnormal state-space modeling. In the first simulation, δ_t was set to be zero for all the observations, which means that the risk factors have no effect across the observations whereas in the second simulation, the known risk factor δ_t to disadvantage the RA-EWMA. However, our simulation doesn't show obvious disadvantage by the risk factor. Through the simulation study, the two methods give very similar results. But, we notice that DGLM is complicated both computationally and conceptually. In contrast, the RA-EWMA estimates by Mean Steady Model is straightforward and simple to implement.

Chapter 3

The Risk-Adjusted EWMA for Poisson Data

The Poisson distribution has a long history of applications to biostatistics problems. It has been traditionally the default distribution to model count data. In this chapter we develop its detailed application under a risk-adjusted EWMA. Our derivations follow closely those done for binary data in Chapter 2 which in turn were described in detail by Grigg and Spiegelhalter (2007). There is no documented detailed analysis for the Poisson case in the literature.

As in Chapter 2, let $d_t = (y_1, ..., y_t)$ denote the data up to time t, which are assumed to be independent Poisson variables,

$$f(y;\mu) = \frac{\mu^y e^{-\mu}}{y!},\ y = 0, 1, 2, ... \tag{3.0.1}$$

The log likelihood function can be written as

$$\log(f(y;\mu)) = y\log(\mu) - \mu - log(y!). \tag{3.0.2}$$

Here the dispersion parameter $\phi = 1$. We will use the log link function, $\eta = g(\mu) = \log(\mu)$.

Adopting a gamma conjugate prior $\pi(\mu_t)$ yields a posterior distribution $p(\mu_t|d_t)$ for the mean parameter μ_t at time t which is also of the gamma form. Following the evolution process, this posterior distribution becomes the prior distribution $\pi(\mu_{t+1}) = p(\mu_t|d_t)$ for the next iteration. After the data y_{t+1} is observed, a Bayesian conjugate update is made, which gives posterior density $p(\mu_{t+1}|d_{t+1})$.

Note that all the distributions involved for this Poisson case have closed forms. If we denote the prior by $p(\mu_t^+|d_t)$ for the mean of an observation with δ_t as the risk-adjustment level, the risk-adjust step can be made by conjugately updating the forecast prior for μ_{t+1}^+ in light of the newly arrived data y_{t+1}.

3.1 The Standard EWMA Model for Poisson Data

Let $p(\mu_t|d_t)$ denote the probability density function of μ_t conditional on the observations up to time t. The pdf is given by the gamma density

$$p(\mu_t; r_t, s_t) = \frac{e^{-s_t\mu_t}\mu^{r_t-1}}{\Gamma(r_t)s_t^{-r_t}}, \quad r_t,\ s_t > 0, \tag{3.1.1}$$

where r_t and s_t are computed from the first t observations. We can write equation (3.1.1) as

$$\mu_t|d_t \sim CP(r_t, s_t). \tag{3.1.2}$$

Consider now the next step and suppose the effect on the distribution is an update of the parameters as follows: $p(\mu_{t+1}|d_t)$ follows a gamma distribution with parameters

r_{t+1} and s_{t+1} such that

$$r_{t+1} = \kappa r_t, \tag{3.1.3}$$

and

$$s_{t+1} = \kappa s_t; \tag{3.1.4}$$

where $0 \leq \kappa \leq 1$. Thus we will have

$$\mu_{t+1}|d_t \sim CP(\kappa r_t, \kappa s_t). \tag{3.1.5}$$

We then have the following properties which are a direct consequence from the gamma distribution

$$\mathbb{E}[\mu_{t+1}|d_t] = \mathbb{E}[\mu_t|d_t] = r_t/s_t, \tag{3.1.6}$$

$$\mathbb{V}[\mu_{t+1}|d_t] = r_t/s_t^2 = \kappa^{-1}\mathbb{V}[\mu_t|d_t], \tag{3.1.7}$$

$$\mathbb{P}[\mu_{t+1}|d_t] = \kappa\mathbb{P}[\mu_t|d_t] = \kappa s_t; \tag{3.1.8}$$

With this specification of the state densities, the parameters μ_{t+1} are related to the best one-step predictor of d_t through the formula

$$\hat{\mu}_{t+1} = \mathbb{E}[\mu_{t+1}|d_t] = r_t/s_t. \tag{3.1.9}$$

The parameters r_t and s_t can be quite arbitrary: Any nonnegative functions of d_t will lead to a consistent specification of the state densities (Brockwell and Davis, 2002).

Once the observation y_{t+1} is available, the posterior distribution $p(\mu_{t+1}|d_{t+1})$ is given by a gamma distribution with parameters $r_{t+1} = \kappa r_t + \phi^{-1}y_t$, and $s_{t+1} = \kappa s_t + \phi$. In Poisson data without considering over-dispersion problem, we simply let $\phi = 1$. We denote this step as

$$\mu_{t+1}|d_{t+1} \sim CP(r_{t+1}, s_{t+1}). \tag{3.1.10}$$

Iterating the relation s_{t+1}, we see that

$$\begin{aligned} s_{t+1} &= 1 + \kappa s_t \\ &= 1 + \kappa + \kappa^2 s_{t-1} \\ &= \cdots \\ &= 1 + \kappa + \kappa^2 + \cdots + \kappa^{t-1} + \kappa^t s_1 \\ &\to 1/(1-\kappa) \end{aligned} \tag{3.1.11}$$

as $t \to \infty$. Similarly,

$$\begin{aligned} r_{t+1} &= y_t + \kappa r_t \\ &= \cdots \\ &= y_t + \kappa y_{t-1} + \cdots + \kappa^{t-1} y_1 + \kappa^t r_1 \end{aligned} \tag{3.1.12}$$

For large t, we have the approximations

$$s_{t+1} = 1/(1-\kappa) \tag{3.1.13}$$

and

$$r_{t+1} = \sum_{i=0}^{t-1} \kappa^i y_{t-i}, \tag{3.1.14}$$

From (3.1.9) the one-step predictors are linear and given by

$$\hat{\mu}_{t+1} = r_{t+1}/s_{t+1} = (\sum_{i=0}^{t-1} \kappa^i y_{t-i})/(\sum_{i=0}^{t-1} \kappa^i). \tag{3.1.15}$$

If we start with $s_0 = 1/(1-\kappa)$, we will find that $\hat{\mu}_{t+1}$ has the following form

$$\hat{\mu}_{t+1} = (1-\kappa) y_t + \kappa \hat{\mu}_t, \quad t = 0, 1, 2, ...,$$

hence by (3.1.5), the one-step predictors can be found by exponential smoothing.

To obtain an ideal one step prediction of $\hat{\mu}_{t+1}$ by EWMA, we need a suitable decay parameter κ. The following steps give the Maximum Likelihood Method to obtain an estimate of κ.

The conjugate prior distribution for Poisson data is a Gamma distribution. Following standard practice, at time $t = 0$ a non informative improper prior is adopted leading to parameter values $r_0 = s_0 = 0$. As a result, a proper posterior distribution for μ_t can be obtained at time $t = \tau$, where y_τ is the first nonzero observation. This is so because since the conjugate prior initial parameter values are $r_0 = s_0 = 0$, thus the parameter updates lead to the same 0 values until a non-zero count comes up for the first time. Thus, for all those cases, the posterior will also be an improper distribution. This in an inherent feature of the method. The problem vanishes if a proper gamma prior is adopted initially.

Now we have, conditional on d_t, the joint density of $y_1, ..., y_T$ is

$$p(y_1, ..., y_T|\kappa) = \prod_{t=\tau+1}^{T} p(y_t|d_{t-1}), \tag{3.1.16}$$

where the predictive probability density functions are given by

$$p(y_t|d_{t-1}) = \int_0^\infty p(y_t|\mu_t)p(\mu_t|d_{t-1})d\mu_t. \tag{3.1.17}$$

In this predictive equation, $p(y_t|\mu_t)$ is from Poisson distribution with pdf

$$p(y_t|\mu_t) = \mu_t^{y_t} e^{-\mu_t}/y_t!, \tag{3.1.18}$$

and $p(\mu_t|d_{t-1})$ follows a gamma distribution with parameters $r_t = \kappa r_{t-1}$ and $s_t = \kappa s_{t-1}$. Actually, we can see that $p(y_t|d_{t-1})$ is a negative binomial distribution, the probability density function can be written as

$$p(y_t|d_{t-1}) = \frac{\Gamma(a + y_t)}{\Gamma(y_t + 1)\Gamma(a)} b^a (1 + b)^{(a+y_t)}, \tag{3.1.19}$$

where $a = \kappa r_{t-1}$ and $b = \kappa s_{t-1}$. Hence the log-likelihood function for unknown parameter κ can be written as

$$l(\kappa) = logL(\kappa) = \sum_{t=\tau+1}^{T} \{\log\Gamma(a+y_t) - \log(y_t!) - \log\Gamma(a) + a\log(b) - (a+y_t)\log(1+b)\}. \tag{3.1.20}$$

The derivative of equation (3.1.20) in terms of unknown parameter κ can be written as following.

$$l'(\kappa) = \sum_{t=\tau+1}^{T} r_{t-1}\{\frac{\gamma(\kappa r_{t-1}+y_t)}{\Gamma(\kappa r_{t-1}+y_t)} - \frac{\gamma(\kappa r_{t-1})}{\Gamma(\kappa r_{t-1})}r_{t-1} - \frac{(\kappa r_{t-1}+y_t)}{(\kappa s_{t-1}+1)}s_{t-1} + \log\frac{\kappa s_{t-1}}{(\kappa s_{t-1}+1)} + 1\}, \tag{3.1.21}$$

where $\gamma(\cdot)$ is digamma function. Although we have not explored the computational aspects in detail, we envision that a Newton-Raphson approach can be used to aproximate the MLE of κ.

3.2 Mean Steady Model for Poisson Data

The mean steady model is a fully parametric model which leads to the RA-EWMA as an estimator for the mean μ_t. For Poisson data, the conjugate prior is a gamma distribution, the link function is a log function, and the forecast prior distribution $\mu_{t+1}^{+}|d_t$ for the mean of an observation with risk-adjustment can be obtained in closed form (Harvey and Fernandes, 1989). In this context, the risk-adjustment step can be made exactly by conjugately updating the forecast prior for μ_{t+1}^{+} based on observation y_{t+1}. For exponential family data, exact risk-adjustment requires that the canonical link function be baseline separable Grigg and Spiegelhalter (2007), the baseline value

can be separated from the risk-factor level either in additive or in multiplicative form. The example for Poisson data is in multiplicative from which can be given by

$$\begin{aligned} \eta_t^+ &= g(\mu_t^+) = g(\mu_t) + \delta_t \\ &\Rightarrow \quad \log(\mu_t^+) = \log(\mu_t) + \delta_t \\ &\Rightarrow \quad \mu_t^+ = e^{\log(\mu_t)+\delta_t} \\ &\Rightarrow \quad \mu_t^+ = \mu_t e^{\delta_t}. \end{aligned} \tag{3.2.1}$$

By Mean Steady Model, the risk-adjusted exponentially weighted moving average method can be written as the following steps.

1. **Conjugate baseline mean.**

$$\mu_t | d_t \sim CP(r_t, s), \tag{3.2.2}$$

where $s = \phi^{-1}(1-\kappa)^{-1}$.

2. **Evolution as Mean Steady Model.**

$$\mu_{t+1} | d_t \sim CP(\kappa r_t, \kappa s). \tag{3.2.3}$$

3. **Update using approximate risk-adjustment.** Using pseudo-observation $\tilde{y}_{t+1}$,

$$\begin{aligned} \mu_{t+1} | d_{t+1} &\sim CP(\kappa r_t + \phi^{-1}\tilde{y}_{t+1}, \kappa s + \phi^{-1}) \\ &\sim CP(r_{t+1}, s). \end{aligned} \tag{3.2.4}$$

where $\tilde{y}_{t+1} = y_{t+1} - (\hat{\mu}_{t+1}^+ - \hat{\mu}_{t+1})$, and $\hat{\mu}_{t+1}^+ = g^{-1}(g(\hat{\mu}_{t+1}) + \delta_{t+1})$. From the above equation, it can be seen that

$$\begin{aligned} \hat{\mu}_{t+1} &= \kappa(r_t/s) + (\phi^{-1}/s)\tilde{y}_{t+1} \\ &= \kappa\hat{\mu}_t + (1-\kappa)\tilde{y}_{t+1}, \end{aligned}$$

which has the same form as equation (2.1.14).

Chapter 4

Application of Poisson Methods to Simulated and Real Data

In this chapter we focus on illustrations of the Poisson methods and case comparisons between the standard and the risk-adjusted EWMA. We follow the approach of Grigg and Spiegelhalter (2007) for the binary case where comparisons were made on cases rather than on extensive simulations to target average performance. With dynamic data, average performance is complicated, particularly in the presence of covariates. The format will be (a) show how to generate Poisson dynamic data, (b) apply the standard and risk-adjusted EWMA methods to the simulated data, (c) apply the methods to a real data set (namely death counts from lung diseases in the UK for the years from 1974 to 1979), and (d) summarize conclusions based on the cases studied.

4.1 Simulated Dynamic Poisson Data

We first present a simulation of dynamic Poisson data with the risk factors set to 0. That is, let $\delta_t = 0,\ t = 1, ..., T$ for all the observations. The conditions been set up are the same as Grigg and Spiegelhalter (2007) for dynamic binary data. Let $\mu_0 = 0.2$, and $W = 0.03$. By this setting, $\hat{v}_t = \mu_0, \phi = 1$ and solving $W \approx \phi(1-\kappa)^2/(\kappa\hat{v}_t)$, we get approximately $\kappa = 0.925$.

In the first stage, η_t, $t = 1, ..., 60$, were simulated for the baseline natural parameter by the distribution $\eta_t|\eta_{t-1} \sim N(\eta_{t-1}, W)$. Second, for $t = 61, ..., 260$, $\eta_t|\eta_{t-1} \sim N(\rho\eta_{t-1}, W)$, where the shift parameter is set to $\rho = 0.99$. Using the simulated data, η_t, to generate μ_t by using the link function $\mu_t = \exp(\eta_t)$. The homogeneous data were generated by sampling from Poisson(μ_t), and the heterogeneous data were generated by sampling from Poisson(μ_t^+). Where $\mu_t^+ = g^{-1}(g(\mu_t) + \delta_t)$. In this study, the risk-adjusted levels were generated from a normal distribution with mean zero and variance 0.25. The initial prior gamma distribution, the distribution of μ_t at time $t = 0$ is noninformative if we choose the initial parameters $r_0 = 0$ and $s_0 = 0$. A proper distribution for μ_t at time $t = \tau$ is obtained when first nonzero observation y_τ arrives. In practise, we usually set the initial value of s_0 other than zero hence set $\tau = 0$.

In Figure 4.1, we compare the estimated value by MSM-EWMA method for the simulated Poisson data with the estimated value by standard EWMA method. We found the performances of these two methods are almost identical except at the very beginning. In this Figure, we give two different initial s_0 for the standard EWMA estimation. Our empirical results show that the estimations with larger initial s_0,

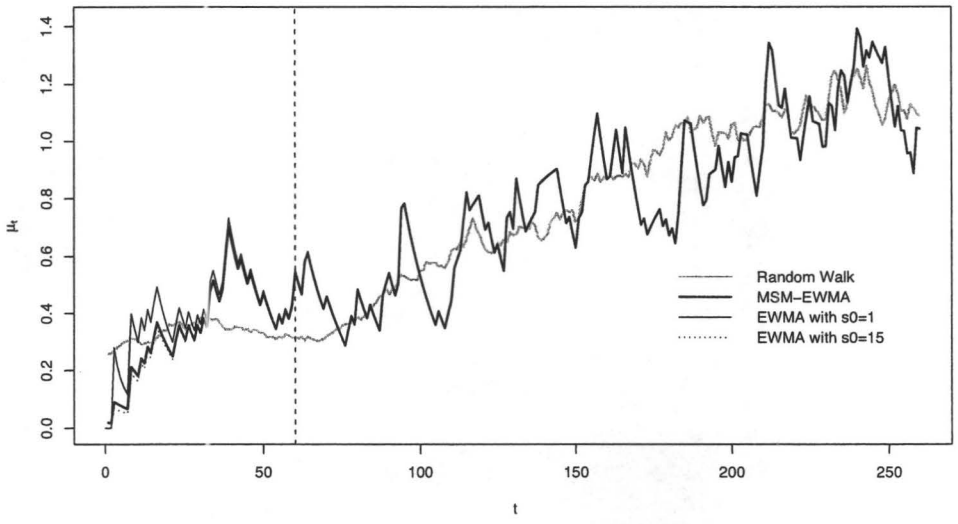

Figure 4.1: *Simulated Poisson homogeneous data. The estimated series by MSM-EWMA and EWMAs methods are displayed whereas the vertical line at* $t = 60$ *marks the random walk starting to shift.*

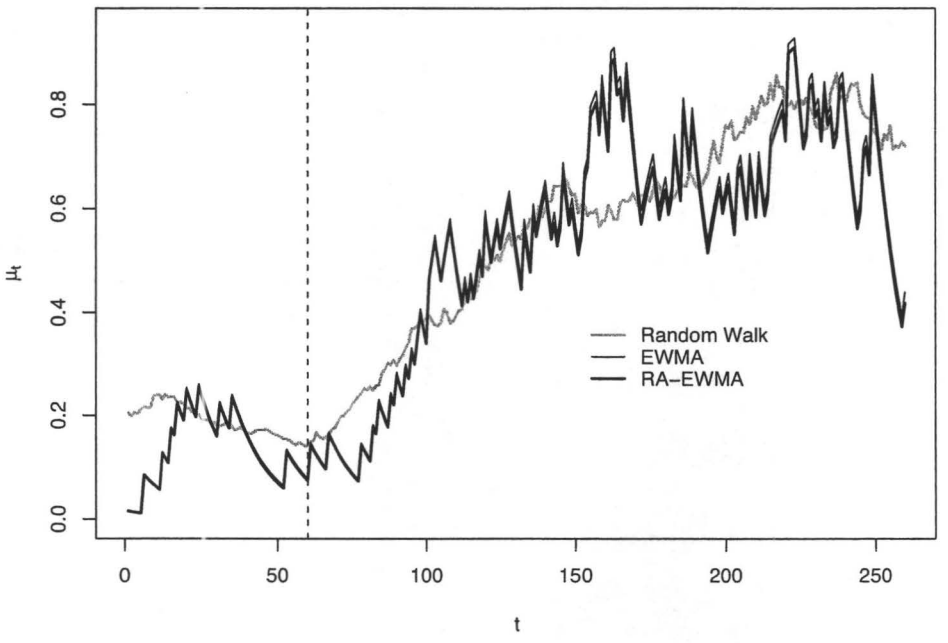

Figure 4.2: *Simulated Poisson heterogeneity data. The estimated series by EWMA and RA-EWMA methods are displayed. The vertical line at* $t = 60$ *marks the random walk starting to shift.*

say $s_0 = 15$, approach to MSM-EWMA much faster than with smaller initial s_0, say $s_0 = 1$.

Figure 4.2 displays the estimated values between the standard EWMA and the risk-adjusted EWMA. We can see that these two estimations are almost identical and both capture the trend exhibited by the random walk. The reason for the similarity in performance might be due to the fact that the risk factor values are relatively small compared to the random walk. However, in some peak points, for example, when t runs from 160 to 170, and from 220 to 240, the estimated values by RA-EWMA method mitigate the effect of the risk factor values. The vertical dashed line marks the change point between the trendless training data and the shifted data.

4.2 Application: Death Counts from Lung Diseases in the UK

The dataset *ldeaths*, which is taken from Diggle (1990), gives the monthly counts of death from bronchitis, emphysema and asthma in the UK for the years from 1974 to 1979. The data were collected monthly and producing a total of 72 observations over the six years. The original data were split by gender. Diggle (1990) gives basic analyses for the data. He noted that the raw data had a seasonal behavior, but when decomposed into seasonal component and residuals, the residuals behaved in a random fashion.

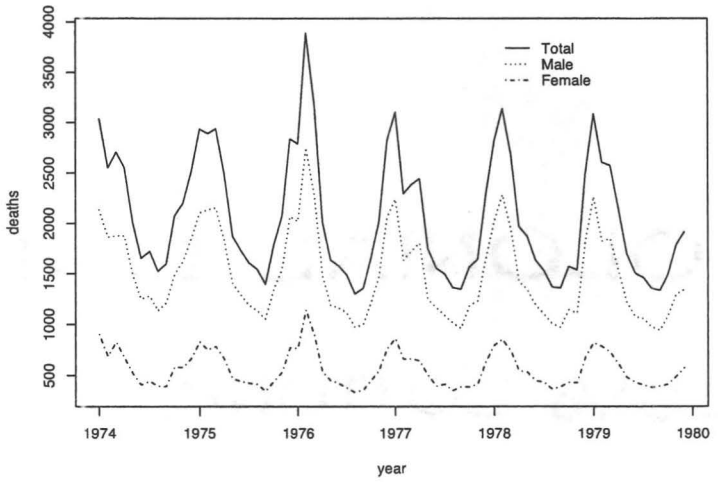

Figure 4.3: *Monthly deaths from lung diseases in the UK for the dataset ldeaths (1974-1979).*

4.2.1 Preliminary Analysis: Seasonal and Residual Components

Figure 4.3 shows the original *ldeaths* data split by gender and the total deaths over the given years. The data clearly show a strong seasonal pattern, with the minimum for each year occurring in July and maximum in February.

Since this is a time series data, we assume that the series X_t runs throughout time, but is observed only for $t = 1, ..., n$, and we denote it by X_t. The series has mean μ. The covariance and correlation functions are given by Venables and Ripley (2002)

$$\gamma_t = cov(X_{t+\tau}, X_\tau), \tag{4.2.1}$$

and

$$\rho_t = \frac{\gamma_t}{\gamma_0} = corr(X_{t+\tau}, X_\tau). \tag{4.2.2}$$

The values of the covariance and correlation function are assumed not to depend on τ, where τ can be any integer.

The second moments are important in the practical analysis of time series since the theory for time series is based on the assumption of second-order stationarity after removing any trends. For $t > 0$ consider the $n-t$ observed pairs $(X_1, X_{1+t}), ..., (X_{n-t}, X_n)$. If we just take the standard correlation or covariance, we use different estimates of mean and variance for each of the subseries $X_{1+t}, ..., X_n$ and $X_1, ..., X_{n-t}$, thus under the second order stationarity assumption, these have the same mean and variance (Brockwell and Davis, 2002). Therefore, we suggest to estimate the autocovariance by Venables and Ripley (2002)

$$c_t = \frac{1}{n}\sum_{s=1}^{n-t}[X_{s+t} - \bar{X}][X_s - \bar{X}], \quad -n < t < n \tag{4.2.3}$$

and estimate the autocorrelation by Venables and Ripley (2002)

$$r_t = \frac{c_t}{c_0}, \quad -n < t < n. \tag{4.2.4}$$

The sequence $\{c_t\}$ and $\{r_t\}$ are the covariance sequence and correlation sequence of the second-order stationary time series. In equation (4.2.3) we use n as denominator even though there are only $n - t$ terms in the sum.

For data containing a trend, the sample autocorrelation function $|\rho(h)|$ will exhibit slow decay as h increases, and for data with a substantial deterministic periodic component, $|\rho(h)|$ will exhibit similar behavior with the same periodicity. Figure 4.4 of the *ldeaths* series shows the seasonal pattern and the autocorrelations do not damp

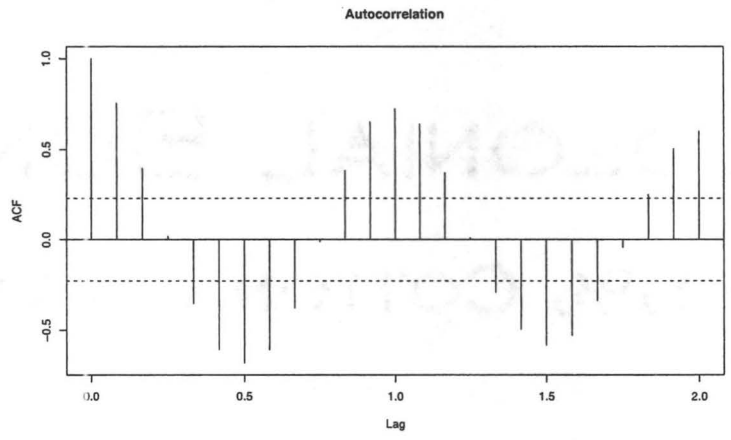

Figure 4.4: *Autocorrelation plot for the ldeaths data.*

down for large lags. In our dataset *ldeaths*, the counts have been collected monthly based. For each twelve months, the data index increased by one. By this property, The lags in Figure 4.4 are expressed in the unit time. For instance, the lag value of 1 indicates one year time unit which includes the data of twelve months. From this autocorrelation plot, we also can see that there is a clear pattern. When the values of lag are at no negative integers such as 0, 1 and so on, they have a locally highest autocorrelation value. The values will gradually down to the lowest after six points, and gradually increase to the highest in next six points. Hence we can say that the data have strongest negative autocorrelation in 6 months period, and strongest positive autocorrelation in 12 months period, which is clearly a seasonal pattern.

The spectral approach to second-order properties is better able to separate short-term and seasonal effects, the detailed theory and formula can be found in Bloomfield (2000). By the theory, the covariance sequence of a second-order stationary time series

can be written as

$$\gamma_t = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega t} dF(\omega) \tag{4.2.5}$$

a finite measure on $(-\pi, \pi]$ for the spectrum F. Under mild conditions that exclude purely periodic components of the series, the measure has a density known as the spectral density f, hence γ_t can be expressed as

$$\gamma_t = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega t} f(\omega) d\omega = \int_{-1/2}^{1/2} e^{2\pi i \omega_f t} f(2\pi\omega_f) d\omega_f. \tag{4.2.6}$$

Where the frequency ω in the first term is in units of radians/time and in the second term ω_f is in unit of cycles/time, where time is in unit of $\triangle t$. If the time series X_t has a frequency greater than one, the spectral density will be divided by frequency.

The Fourier integral can be inverted as

$$f(\omega) = \sum_{-\infty}^{\infty} \gamma_t e^{-i\omega t} = \gamma_0[1 + 2\sum_{1}^{\infty} \rho_t \cos(\omega t)]. \tag{4.2.7}$$

By the symmetry of γ_t, $f(\omega) = f(-\omega)$, we need only consider f on $(0, \pi)$. A smoother estimate of ω can also be derived. The periodogram is related to the autocovariance function by

$$I(\omega) = \sum_{-\infty}^{\infty} c_t e^{-i\omega t} = c_0[1 + 2\sum_{1}^{\infty} r_t \cos(\omega t)] \tag{4.2.8}$$

and

$$c_t = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega t} I(\omega) d\omega. \tag{4.2.9}$$

We omit the details which are given by Bloomfield (2000) and Brockwell and Davis (2002). Figure 4.5 gives spectral density and cumulative periodogram estimates for the *ldeaths* data. The bandwidth in the spectral density plot is a measure of the size of the smoothing window. If there are periodic components in the series, there will

be peaks in the spectral density plot. Clearly, there is a peak at frequency equal to 1, which indicates the data have periodic component with period 1. In regards to the *ldeaths* data, we conclude that the data have a one year period. In the spectral density plot, we use a smoother which equals to 3. The smoothing uses the modified Daniell smoothers Bloomfield (2000), which are moving averages giving half weight to the end values of the span. The smoothing will reduce those peaks, but they can be seen quite clearly in the plot of the cumulative periodogram, the two dashed lines display the 95% confidence band for the cumulative plot.

From above analyses, we learned that the data have a one year (twelve months) period and a seasonal effect. The next question is how to decompose the raw data. In Brockwell and Davis (2002), the classical decomposition model been given by

$$X_t = m_t + s_t + Y_t, \tag{4.2.10}$$

where m_t is a slowly changing function known as a trend component, s_t is a function with known period d named as seasonal component, and Y_t is a stationary random noise, called residual component. Cleveland *et al.* (1990) proposed a method to detrend a time series using Local Polynomial Regression Fitting. The basic idea for this method can be described as follows. Since we already found the period is one year, we collect the sub-series, say all the data for January, February, as well as other months, smoothing them by replacing the data with the mean which gives the seasonal component. After the seasonal values been removed, we then smooth the remainder to find the trend. This leads to a decomposition of the raw data into three parts: seasonal component, remainder component, and trend component.

Figure 4.6 shows the decomposition of the raw data. Which includes the Original

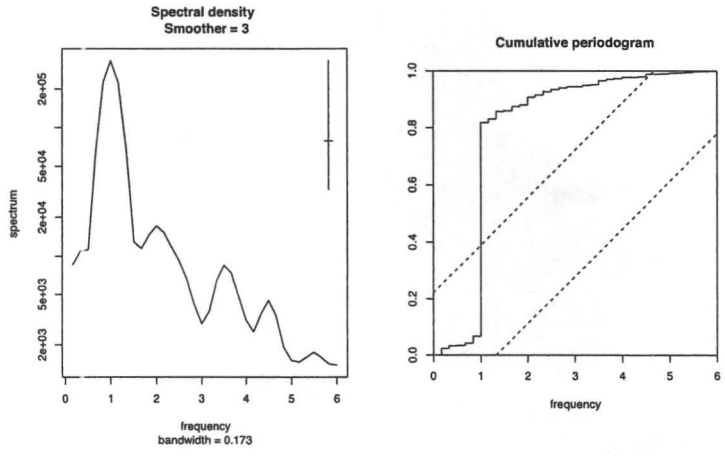

Figure 4.5: *Spectral density estimates and Cumulative periodogram for the ldeaths data.*

data, seasonal component, remainder component, and the trend component.

In next section, we will apply the EWMA and RA-EWMA methods to the residual component, and compare the results for different conditions.

4.2.2 The EWMA and RA-EWMA Methods for Lung Disease Death Data

In this section, we apply the standard EWMA and RA-EWMA methods to the total deaths in the *ldeaths* data collected over 1974-1979. Two comparisons are pursued. First, the comparison between the non-seasonal component, trend component, and standard EWMA estimate applied to the non-seasonal component. Second, the comparison between the original data, the standard EWMA estimation on the original data, and the risk-adjusted EWMA using the transformation of seasonal component

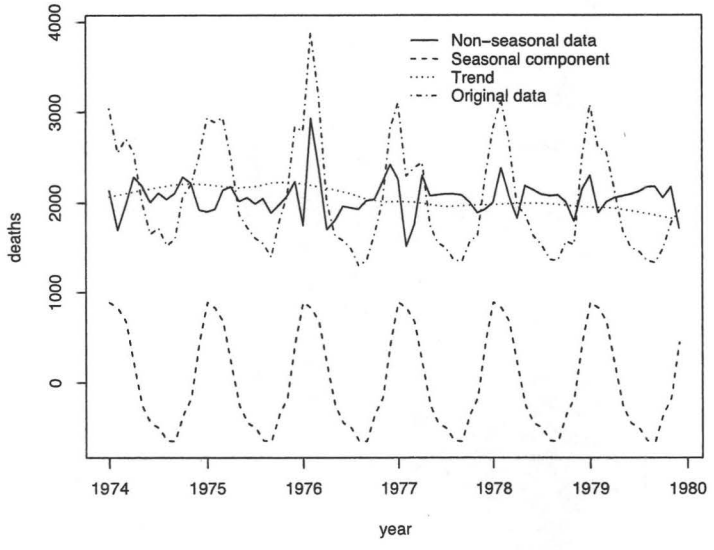

Figure 4.6: *The decomposition for the ldeaths data.*

as risk factor.

We first consider a standard exponentially weighted moving average analysis by assuming that the non-seasonal data are Poisson distributed with mean μ_t and estimate μ_t by a standard EWMA with decay parameter κ; then compare this EWMA with the trend component and the non-seasonal data. Since we do not have historical data to estimate κ, we set $\kappa = 0.925$, this value is not uncommon in the use of the standard EWMA. From equation (2.2.5), we have $s = \phi^{-1}(1-\kappa)^{-1}$. Recall that for the Poisson data, the dispersion parameter ϕ is equal to 1. Treating the EWMA as the limiting posterior mean of an MSM without risk-adjustment, the posterior distribution for μ_t after each observation is taken to be a gamma distribution with parameters $r_{t+1} = r_t + y_t$, and $s_{t+1} = s_t + 1$. The initial value of r_0 can be set as zero as discussed before.

Figure 4.7 displays the relevant plots. Comparing the EWMA with the trend data, it is clear that the EWMA estimate follows the hidden trend of the Deaths Data.

Assuming now that the original data are Poisson(μ_t^+), we can dynamically estimate μ_t by a RA-EWMA, where risk factors are assumed to affect the outcomes. In this example, we use a log linear regression to estimate the seasonal component as the risk factor. By equation (3.2.1), we have the relation $\mu_t^+ = \mu_t e^{\delta_t}$, where δ_t is the risk factor.

Figure 4.8 gives the EWMA control chart for non-seasonal component with trend. The dots represent the values of the non seasonal component. There are two glaring features shown this plot. First, the EWMA estimate plot is close to the trend. Second, the EWMA estimates have few points out of control around years 1976 and 1977, which were caused by several usually large observations. At the beginning of year

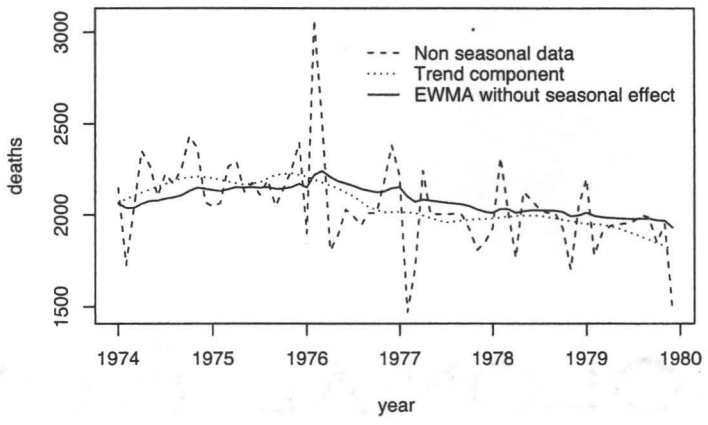

Figure 4.7: *The estimates series by the standard EWMA method for the ldeaths data without seasonal component, non-seasonal component, and trend component.*

1977, there are also very low observed values, but the EWMA estimates are still in control. This is due to the fact that there exist several large observations before these particular small observations. Hence influences the estimation after those high values.

In the last analysis, we applied the standard Exponentially Weighted Moving Average and the risk-adjusted Exponentially Weighted Moving Average methods to the residuals from the *ldeaths* data, without seasonal component and trend component. In Figure 4.9, the EWMA control chart for non-seasonal data without trend component (i.e. the residuals obtained by removing the seasonal and trend components) is displayed. The main message from this plot is that the EWMA chart is doing very well with nearly all the EWMA estimate values falling in the in-control region except for one point close to the upper limit, which is caused by the one high original observation.

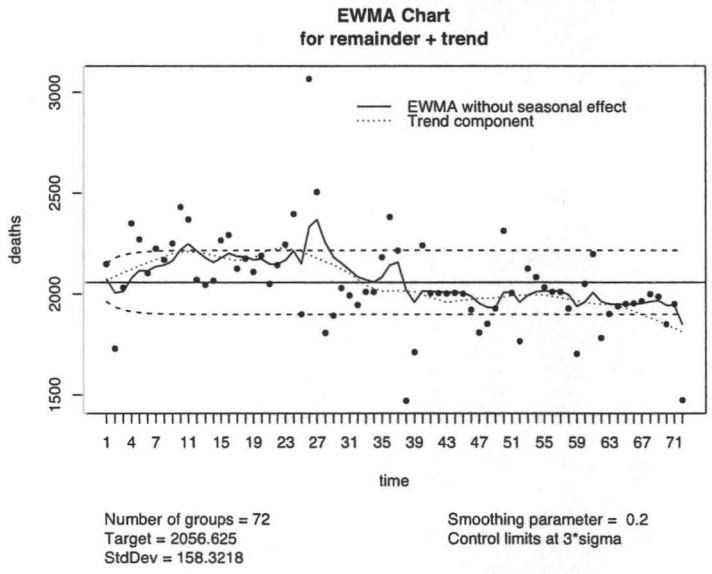

Figure 4.8: *The estimates series by the standard EWMA method for the ldeaths data without seasonal component, with 3σ as the control limit.*

Figure 4.9 shows that the RA-EWMA is more sensitive to the data than the EWMA method. As the original data change, the effect is shown on RA-EWMA immediately whereas for the standard EWMA, the effect was masked by the previous observations. From a healthcare monitoring perspective, the sensitivity of the RA-EWMA method can help healthcare providers make adjustments to the services and treatments provided, avoiding unnecessary delays. However, one outlier outcome might cause the RA-EWMA based monitoring system to sound alarm. This may or may not be a good thing to happen, depending on the true cause of the outlier. For instance, if it was due to the fact that something went wrong with that particularly patient, then the alarm would be justified. Note, however, that because the RA-EWMA and the standard EWMA are moderated by previous observations, it will not happen in all cases that an outlier will automatically cause the issuing of an alarm.

Table 4.1 gives the numerical estimates by the standard EWMA, the RA-EWMA methods for the remainder component, namely the residuals, in the *ldeaths* data in the year of 1976. The second column of the table is the remainder component, for year 1976 after the decomposition for the original data. The calculated mean and standard deviation of the differences between the residuals and the estimated values ($\hat{\mu}_t^E$) by the standard EWMA method are given by 26.33 and 334.46 respectively. Similarly, the mean and standard deviation of the difference between the residuals and the estimated values ($\hat{\mu}_t^R$) by the RA-EWMA method are given by 37.08 and 343.83 respectively. The latter gives larger mean and standard deviation. The results show that EWMA gives a smooth change, but RA-EWMA is more sensitive due to

Table 4.1: *The numerical estimates by the standard EWMA and the RA-EWMA methods for remainder component, namely residuals, in the ldeaths data in the year of 1976.*

Month	Residual	$\hat{\mu}_t^R$	$\hat{\mu}_t^E$
01	1747	2067	2034
02	2930	2020	2101
03	2385	2163	2122
04	1702	1969	2091
05	1804	2089	2069
06	1961	2028	2061
07	1946	2056	2053
08	1928	2048	2043
09	2024	2047	2042
10	2038	2054	2041
11	2224	2056	2055
12	2422	2069	2083

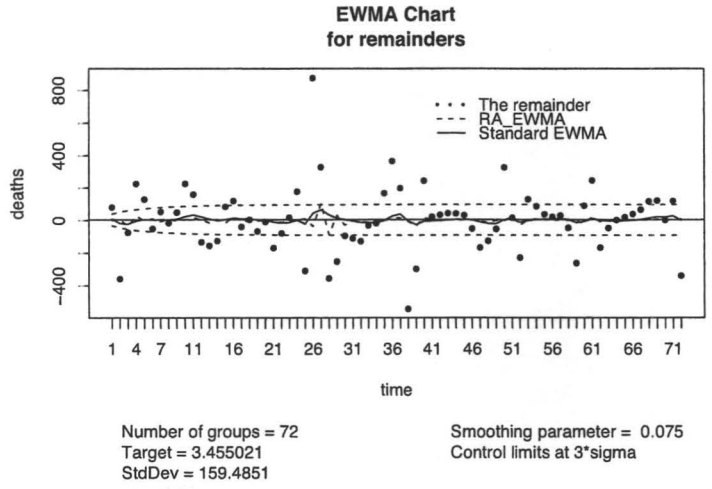

Figure 4.9: *The estimates series by the standard EWMA and the RA-EWMA methods for remainder component, with 3σ as the control limit, for the ldeaths data.*

the risk-adjustment.

Chapter 5

Conclusions and Future Studies

In this thesis, we developed in detail the Risk-Adjusted Exponentially Weighted Moving Average for Poisson data, a method useful to monitor healthcare or other counts that are generated dynamically over time. The approach is motivated by and follows the approach used by Grigg and Spiegelhalter (2007) for dynamic binary outcomes. The basic idea is to use a Bayesian approach with conjugate priors (gamma in the case of Poisson data) that are used iteratively to provide the estimates. The main application is the counts with covariates. The necessary formulas are given to update the estimates from the method.

The numerical evidence provided in the thesis suggests that the RA-EWMA method is more sensitive than the standard EWMA method to the presence of the underlying covariates. This was shown clearly on the real data, specifically in the UK's death counts from lung diseases.

When we develop our formulas for both EWMA and RA-EWMA methods for Poisson data, we simply assume the dispersion parameter, namely ϕ, as 1. But

in the most applications for Poisson data, the over-dispersion problem needs to be considered. The initial thinking can be addressed as following. In the monitoring process, we usually use test data or training data to estimate the initial values such as r_0, s_0, and κ. At the same time, we also can use the test data or training data to estimate the value of the dispersion parameter. Once we have the the estimated dispersion parameter, we can update equation (3.1.10) by $r_{t+1} = \kappa r_t + \phi^{-1} y_t$, and $s_{t+1} = \kappa s_t + \phi$ instead of by $r_{t+1} = \kappa r_t + y_t$, and $s_{t+1} = \kappa s_t + 1$.

In this thesis, we take β and other parameters such as σ^2 to be known. In practice, however, the process parameters are estimated from data gathered when the process was operating in control. This is the approach followed in nearly all the quality control applications. In real healthcase process, the risk factor might change over time due to the reassessment of the patients. For example, when a patient is reassessed in the quarter assessment or full assessment, there may be some new covariates which become significant hence the old risk factor coefficients need to be updated. In the monitoring process, it is recommended that one periodically reassesses the risk model. In our application, we simply use the transformation of seasonal information as risk factor, this might be inadequate for a real process. Hence the method may not pick effectively the hidden information in the raw data.

In real healthcare monitoring processes, there are many possible applications and further research problems. One important issue is the derivation of appropriate control limits to signal out-of-control excursions of the healthcare process. In the thesis we used the conventional $\pm 3\sigma$, but quite likely better control limits should be used for the RA-EWMA. In order to calibrate the method for a particular healthcare process, one needs to compute the associated average run lengths to guide practitioners.

So far, only discrete outcomes (i.e. binary or Poisson) have been discussed in detail. But there are many healthcare processes where continuous outcomes are being monitored, for instance the remission times of patients receiving a particular treatment. A log-normal or gamma distribution for the data may be more appropriate. Nothing has been done in this direction.

Appendix A

Notation Index

Symbol	Description	Page
CQI	continuous quality improvement	1
CUSUM	cumulative sum	3
DGLM	dynamic generalized linear model	19
EWMA	exponentially weighted moving average	x
GLM	generalized linear model	9
MLE	maximum likelihood estimation	23
MSM	mean steady model	18
RA-EWMA	risk-adjusted exponentially weighted moving average	x
SPC	statistical process control	1
TQM	total quality management	1

Appendix B

R Program for Simulation Studies and Application

B.1 Simulation Study for Bernoulli Data

```
###============================================================###
###                                                            ###
### Review for the Bernoulli data example                      ###
### Mortality rate after Cardiac surgery                       ###
###                                                            ###
### Main referenced Paper                                      ###
### A simple risk-adjusted exponentially weighted moving average ###
### Olivia Grigg and David Spiegelhalter                       ###
### Journal of the American Statistical Association, March 2007  ###
###                                                            ###
###============================================================###

# The R code need to use library Rlab

# Rlab is a collection of functions and datasets to be used in
# the class ST370-Probability and Statistics for Engineers at
# North Carolina State University.

library(Rlab)

# For more information see the class labs at:
# http://www.courses.ncsu.edu/ST370/distance/rlab/rlab.html
```

```
# These labs are based on Slab and Mlab
# by Doug Nychka and Dennis Boos.


# The following R code has been used to
# plot Figure 2.1
# the simple sample for the surgery data
 ra.ewma <- c(.088,.082,.149,.113,.137)
 std.ewma <- c(.158,.142,.228,.205,.285)
 hat.ewma <-c(.596,.058,.328,.359,.764)

# Figure 2.1: EWMA and RA-EWMA for Cardiac surgery data
 y <- c(1,0,1,0,1)
 plot( y, type='p', pch=20, xlab='time', ylab='estimation')

 points(ra.ewma,type='b',  lty=1)
 points(std.ewma,type='b', lty=2)
 points(hat.ewma,type='b', lty=3)

 msg <- c('Risk adjusted EWMA', 'Standard EWMA',
           'Estimate without risk adjustment')
 legend(2,0.9, lty=c(1,2,3),msg, bty='n')



# Formulas reference this thesis Section 2.4.
# Simulation study for Binary Data

# Random-walk
# To generate random data for mu_t's, to compare
# the estimated series by DGLM and EWMA

# If want get the same simulation data
# we can use seed()

# set n1 for the number of trendless training data
# set n2 for the number of shifted data
n1 <- 60
n2 <-200
n <- n1 + n2

# w is for the variance to generate random data
w <- 0.03
# given the initial value of mu0
mu0 <- 0.2
eta0 <- log(mu0/(1 - mu0))
# set.seed to generate same sample
#set.seed(1)
mu <- eta <- NULL
```

```
eta[1] <- rnorm(1,eta0,w)

# Loop for generating trendless training data
for (i in 2:n1){
    #set.seed(i)
    eta[i] <- rnorm(1,eta[i-1],w)
}

# shift parameter
rho <- 0.01

# Loop for generating shifted data
for (i in (n1+1):(n)){
    #set.seed(i)
    eta[i] <- rnorm(1,(1-rho) * eta[i-1],w)
}

# inversed link function for Binary data
mu <- exp(eta)/(1 + exp(eta))


# DGLM
# Formulas reference this thesis Section 2.3.1.
# There are detailed description for each step.
# the repetition to record estimated series
# average to get the final result
# the amount depend on the required accuracy
mm<-500

# The Homogeneity Bernoulli data
yho <- rbern(n, mu)

# Risk asjustment level at zeros
de <- rnorm(n*mm, 0, 0)
delta <- array(data=de, dim=c(260,mm))


# following code to do the DGLM estimation
# the step number following the process number
# in section 2.3.1:
# Dynamic Generalized Linear Model

fs <- fr <- array(data = 0, dim = c(260,mm))
for (j in 1:mm ){

    etaplus  <- eta + delta[,j]
    muplus <- exp(etaplus)/(1 + exp(etaplus))

    m <- NULL
    C <- NULL
```

```
    r <- s <- NULL
    rrt <-sst <- NULL


    sst[1] <- s[1] <-1

    rrt[1] <- r[1] <- mu[1]

    for (t in 1:(n-1)){
        # step 2
        m[t] <- log(r[t]/(s[t] - r[t]))
        C[t] <- 1/r[t] + 1/(s[t] - r[t])

        # step 4
        et <- m[t] + delta[t+1,j]
        qt <- C[t] + w

        # step 5
        rrt[t+1] <- (1 + exp(et))/qt
        sst[t+1] <- (2 + exp(et)+ exp(-et))/qt

        # step 7
        rrrt <- rrt[t+1] + yho[t+1]
        ssst <- sst[t+1] +1

        eet <- log(rrrt/(ssst - rrrt))
        qqt <- 1/rrrt + 1/(ssst - rrrt)

        m[t+1] <- eet - delta[t+1,j]
        C[t+1] <- qqt

        # step 8
        r[t + 1] <- (1 + exp(m[t+1]))/C[t+1]
        s[t +1] <- (2 + exp(m[t+1])+ exp(-m[t+1]))/C[t+1]

    }

    fs[,j] <- sst
    fr[,j] <- rrt

}

# average the simulated estimate values
# to get the estimated parameter r_t and s_t
# for all t.
mfs <- rowMeans(fs, dims = 1)
mfr <-  rowMeans(fr, dims = 1)


# Estimation by Mean Steady Model
```

```
# For bernoulli data phi = 1
phi <- 1
kappa <- 0.93          # for mu0=0.2

rar <- NULL
rar[1] <- mu[1]
ras <- phi^(-1) * (1 - kappa)^(-1)
rafr <- array(data = 0, dim = c(n,mm))

for (j in 1:mm ){

    for (t in 1:(n-1)){

        # step 1,2
        muhat <- rar[t]/ras

        # step 3
        g <- log(muhat /(1 - muhat))
        ginv <- exp(g + delta[t+1,j])
                /(1 + exp(g + delta[t+1,j]))
        ytilde <- yho[t+1] - (ginv - muhat)

        # step 3, update
        rar[t + 1] <- kappa * rar[t] + phi^(-1)* ytilde

    }

    rafr[,j] <- rar/ras

}

mrafr <-  rowMeans(rafr, dims = 1)
ymax <- max(mu,(mfr/mfs),mrafr)

# calculate the differencial risk
drisk <- exp(etaplus)/(1 +exp(etaplus))

# the following code been used to plot Figure 2.2
# Figure 2.2: simulation RA-EWMA and DGLM for Bernoulli
# homogeneous data

plot(2:n, mu[-1], type='l',xlab = 't',
        ylab=expression(mu[t]),
        ylim=c(0,ymax + 0.02),lwd=2, col=8)
abline(v=n1,lty=2)
points(2:n,mfr[-1]/mfs[-1], type='l',lwd=2)
points(2:n,mrafr[-n], type='l',lwd=1)
points(2:n,drisk[-1], type='p',pch='.',cex=2)

msg <- c('Random Walk','DGLM','EWMA','Differential Risk')
```

```
legend(200,0.2, msg, lty=c(1,1,1,3),lwd=c(2,2,1,2),
       col=c(8,1,1,1),bty='n')

#=================================================
# the following code been used to calculate the
# estimated values for Figure 2.3
# The code is same as that for Figure 2.1
# except this is for heterogeneous data

# the DGLM estimation
# Formulas reference this thesis Section 2.3.1.
# There are detailed description for each step.

# delta is the risk-adjustment level
# generated from normal distribution
# with mean 0 and variance 0.25
var = 0.25
sds= sqrt(var)
delta <- rnorm(n, 0, sds)
etaplus  <- eta + delta
muplus <- exp(etaplus)/(1 + exp(etaplus))
yhe <- rbern(n, muplus)  ## The Heterogeneity data

de <- rnorm(n*mm, 0, sds)
# array delta for the risk-adjustment level
delta <- array(data=de, dim=c(260,mm))

# following code to do the DGLM estimation
# the step number following the process number
# in section 2.3.1:
# Dynamic Generalized Linear Model

# the DGLM estimation srarted
fs <- fr <- array(data = 0, dim = c(260,mm))
for (j in 1:mm ){

    etaplus  <- eta + delta[,j]
    muplus <- exp(etaplus)/(1 + exp(etaplus))

    m <- NULL
    C <- NULL
    r <- s <- NULL
    rrt <-sst <- NULL


    sst[1] <- s[1] <-1

    rrt[1] <- r[1] <- mu[1]

    for (t in 1:(n-1)){
```

```
        # step 2
        m[t] <- log(r[t]/(s[t] - r[t]))
        C[t] <- 1/r[t] + 1/(s[t] - r[t])

        # step 4
        et <- m[t] + delta[t+1,j]
        qt <- C[t] + w

        # step 5
        rrt[t+1] <- (1 + exp(et))/qt
        sst[t+1] <- (2 + exp(et)+ exp(-et))/qt

        # step 7
        rrrt <- rrt[t+1] + yhe[t+1]
        ssst <- sst[t+1] +1

        eet <- log(rrrt/(ssst - rrrt))
        qqt <- 1/rrrt + 1/(ssst - rrrt)

        m[t+1] <- eet - delta[t+1,j]
        C[t+1] <- qqt

        # step 8
        r[t + 1] <- (1 + exp(m[t+1]))/C[t+1]
        s[t +1] <- (2 + exp(m[t+1])+ exp(-m[t+1]))/C[t+1]

    }

    fs[,j] <- sst
    fr[,j] <- rrt

}
# calculate the mean of mm times estimated values
mfs <- rowMeans(fs, dims = 1)
mfr <-  rowMeans(fr, dims = 1)


# the RA-EWMA estimation
# Use MSM method
# DGLM
# Formulas reference this thesis Section 2.3.2
# There are detailed description for each step.

# For bernoulli data phi = 1
phi <- 1
kappa <- 0.93           # for mu0=0.2

rar <- NULL
rar[1] <- mu[1]
ras <- phi^(-1) * (1 - kappa)^(-1)
```

```
rafr <- array(data = 0, dim = c(260,mm))

for (j in 1:mm ){

    for (t in 1:(n-1)){

        # step 1,2
        muhat <- rar[t]/ras

        # step 3
        g <- log(muhat /(1 - muhat))
        ginv <- exp(g + delta[t+1,j])
                /(1 + exp(g + delta[t+1,j]))
        ytilde <- yhe[t+1] - (ginv - muhat)

        # step 3, update
        rar[t + 1] <- kappa * rar[t] + phi^(-1)* ytilde
    }
    rafr[,j] <- rar/ras
}

# calculate the mean of the estimations
mrafr <-  rowMeans(rafr, dims = 1)

# calculate the ylimit for the plot
ymax <- max(mu,(mfr/mfs),mrafr)

# calculate the differential risk
drisk <- exp(etaplus)/(1 +exp(etaplus))

# Figure 2.3: Simulation RE-EWMA and DGLM for Bernoulli
# heterogeneous data
plot(2:n, mu[-1], type='l',xlab = 't',
        ylab=expression(mu[t]),
        ylim=c(0,ymax + 0.02),lwd=2, col=8)
abline(v=n1,lty=2)
points(2:n,mfr[-1]/mfs[-1], type='l',lwd=2)
points(2:n,mrafr[-n], type='l',lwd=1)
points(2:n,drisk[-1], type='p',pch='.',cex=2)

msg <- c('Random Walk','EWMA','DGLM','Differential Risk')
legend(200,0.2, msg, lty=c(1,1,1,3),
        lwd=c(2,2,1,2),col=c(8,1,1,1),bty='n')
```

B.2 Simulation Study for Poisson Data

```
###================================================================###
###                                                                ###
### Generate Poisson data and programming                          ###
###                                                                ###
### Main referenced Paper                                          ###
### A simple risk-adjusted exponentially weighted moving average   ###
### Olivia Grigg and David Spiegelhalter                           ###
### Journal of the American Statistical Association, March 2007    ###
###                                                                ###
### Another referenced Paper                                       ###
### Time series models for Count or Qualitative Ovservations       ###
### A.C. Harvey and C. Fernandes                                   ###
### Journal of Business & Economic Statistics, October 1989        ###
###================================================================###

# Generate Random-walk data for Poisson data
# If want get the same simulation data
# we can use seed()

# set n1 for the number of trendless training data
# set n2 for the number of shifted data

n1 <- 60
n2 <-200
n <- n1 + n2
w <- 0.03
mu0 <- 0.2
eta0 <- log(mu0)
#set.seed(1)
mu <- eta <- NULL
eta[1] <- rnorm(1,eta0,w)
for (i in 2:n1){
    #set.seed(i)
    eta[i] <- rnorm(1,eta[i-1],w)
}

rho <- 0.01
for (i in (n1+1):(n)){
    #set.seed(i)
    eta[i] <- rnorm(1,(1-rho) * eta[i-1],w)
    #eta[i] <- rnorm(1,rho * eta[i-1],w)
}

mu <- exp(eta)

#================================================
# the repetition to record estimated series
```

```
# average to get the final result
# the amount depend on the required accuracy

# the repetition mm set as 500
mm<-500

delta <- rnorm(n, 0, 0.25)
etaplus = eta + delta
muplus <- mu * exp(delta)

yhe <- rpois(n, muplus)
# The Heterogeneity poisson data
# Baseline mean plus risk adjustment level
# Risk asjustment level at 0.25
de <- rnorm(n*mm, 0, 0.25)

# risk adjust levels
delta <- array(data=de, dim=c(260,mm))

# risk adjust levels are zeros for
# homogeneous data
delta0 <- array(data=0, dim=c(260))

# Possion data, dispersion parameter is equals to one
phi <- 1

# selected decay parameter
kappa <- 0.925

#=================================================
# EWMA without risk adjustment

# The formulas please reference section 3.2

# The following paper has some discuss about possion case
# Harvey, A., and Fernandes, C. (1989)
# Time Series Models for Count or Qualitative Observations
# Journal of Business and Economic Statistics, 7, 407-417.
# Section 2


a <- NULL
b <- NULL

a0 <- 0
b0 <- 1

# the following code for Figure 4.1
# the comparison between different initial
# value of s0 for RA-EWMA method, and MSM method
```

```
# the EWMA with initial s0=1
muhat18 <- array(data = 0, dim = c(260))

# the predicted values of mu by MSM-EWMA

    for (t in 1:(n-1)){

        # step 1
        if (t==1){
            at <- kappa * a0
            bt <- kappa * b0
        }
        else {
            at <- kappa * a[t-1]
            bt <- kappa * b[t-1]
        }
        muhat18[t] <- at/bt

        # step 2, update
        a[t] <- at + yhe[t]
        b[t] <- bt + 1
    }

mpois1 <- muhat18

# the following code for Figure 4.1
# the EWMA with initial s0=15
a0 <- 0
b0 <- 15

muhat18 <- array(data = 0, dim = c(260))

# the predicted values of mu by MSM-EWMA

    for (t in 1:(n-1)){

        # step 1
        if (t==1){
            at <- kappa * a0
            bt <- kappa * b0
        }
        else {
            at <- kappa * a[t-1]
            bt <- kappa * b[t-1]
        }
        muhat18[t] <- at/bt

        # step 2, update
        a[t] <- at + yhe[t]
```

```
        b[t] <- bt + 1
    }

mpois15 <- muhat18

# EWMA without risk adjustment
# MSM-EWMA in Figure 4.1
# For Poisson data phi = 1,
phi <- 1
kappa <- 0.925

rar0 <- NULL
rar0[1] <- mu[1]
ras0 <- phi^(-1) * (1 - kappa)^(-1)

# define an array to save the one step prediction values
rafr0 <- array(data = 0, dim = 260 )

for (t in 1:(n-1)){

    # step 1,2
    muhat0 <- rar0[t]/ras0

    # step 3
    g <- log(muhat0)
    ginv0 <- muhat0
    ytilde0 <- yhe[t+1] - (ginv0 - muhat0)

    # step 3, update
    rar0[t + 1] <- kappa * rar0[t] + phi^(-1)* ytilde0
}

rafr0 <- rar0/ras0

# calculate the y limit for the plot
ymax <- max(mu,mpois1,mpois15, rafr0)

# Plot to compare the EWMA with different initial values b,
# and EWMA in Main paper.  without risk adjustment

# Figure 4.1 EWMA with different initial values
plot(1:n, mu, type='l',xlab = 't', ylab=expression(mu[t]),
       ylim=c(0,ymax + 0.02),lwd=2, col=8)
abline(v=n1,lty=2)
points(1:n,rafr0, type='l',lwd=2)   # EWMA

points(0:(n-2),mpois1[-n], type='l',lty=2,lwd=1)
points(0:(n-2),mpois15[-n], type='l',lty=3,lwd=1)
```

```
msg <- c('Random Walk','MSM-EWMA','EWMA with s0=1',
    'EWMA with s0=15')
legend(200,0.4, msg, lty=c(1,1,2,3),lwd=c(2,2,1,1),
        col=c(8,1,1,1),bty='n')

#=================================================
# the following code for Figure 4.2
# EWMA without risk adjustment
# For Poisson data phi = 1,
phi <- 1
kappa <- 0.925          # for mu0=0.2

rar0 <- NULL
rar0[1] <- mu[1]
ras0 <- phi^(-1) * (1 - kappa)^(-1)
rafr0 <- array(data = 0, dim = 260 )

for (t in 1:(n-1)){

    # step 1,2
    muhat0 <- rar0[t]/ras0

    # step 3
    g <- log(muhat0)
    ginv0 <- muhat0
    ytilde0 <- yhe[t+1] - (ginv0 - muhat0)

    # step 3, update
    rar0[t + 1] <- kappa * rar0[t] + phi^(-1)* ytilde0
}

rafr0 <- rar0/ras0

# the following code for Figure 4.1
# RA-EWMA, with risk adjustment
# For Poisson data phi = 1,
phi <- 1

rar <- NULL
rar[1] <- mean(yhe)
ras <- phi^(-1) * (1 - kappa)^(-1)
rafr <- array(data = 0, dim = c(260,mm))

for (j in 1:mm ){

    for (t in 1:(n-1)){

        # step 1,2
        muhat <- rar[t]/ras
```

```
        # step 3
        g <- log(muhat)
        ginv <- muhat * exp(delta[t+1,j])
        ytilde <- yhe[t+1] - (ginv - muhat)

        # step 3, update
        rar[t + 1] <- kappa * rar[t] + phi^(-1)* ytilde
    }

    rafr[,j] <- rar/ras
}

mrafr <-  rowMeans(rafr, dims = 1)
ymax <- max(mu,mrafr, rafr0)

drisk <- exp(etaplus)

# Figure 4.2: Comparison between EWMA and RA-EWMA
plot(1:n, mu, type='l',xlab = 't', ylab=expression(mu[t]),
        ylim=c(0,ymax + 0.02),lwd=2, col=8)
abline(v=n1,lty=2)
points(1:n,rafr0, type='l',lwd=1)        # EWMA
points(1:n,mrafr, type='l',lwd=2)        # RA-EWMA

msg <- c('Random Walk','EWMA','RA-EWMA')
legend(150,0.4, msg, lty=c(1,1,1),
        lwd=c(2,1,2),col=c(8,1,1),bty='n')

# Figure 4.3: Control Chart --
#Comparison between EWMA and RA-EWMA
q <- qcc(yhe[1:60] , newdata=yhe[61:150],
    type="xbar.one", plot=FALSE)
qcc.options(bg.margin="white")
ewma(q,lambda = (1-kappa),xlab='time',ylab='Poisson data')
points(1:150,mrafr[1:150],type='l',lty=2)
msg <- c('Standard EWMA','RA-EWMA')
legend(110,1.75,lty=c(1,2),msg, bty='n')
```

B.3 Application for Lung Deaths Data

```
###============================================================###
###                                                            ###
### Application for Count data                                 ###
### Monthly Deaths from Lung Diseases in the UK                ###
### Main reference paper:                                      ###
### A simple risk-adjusted exponentially weighted moving average ###
### Olivia Grigg and David Spiegelhalter                       ###
### Journal of the American Statistical Association, March 2007  ###
###                                                            ###
### Another referenced Paper                                   ###
### Time series models for Count or Qualitative Ovservations   ###
### A.C. Harvey and C. Fernandes                               ###
### Journal of Business & Economic Statistics, October 1989    ###
###                                                            ###
### The ldeaths data is from:                                  ###
### P. J. Diggle (1990) Time Series:                           ###
### A Biostatistical Introduction. Oxford, table A.3           ###
### It now included in R package stats library                 ###
###                                                            ###
###============================================================###


# Step 1: Decompose the original data to seasonal,
# trend, and remainder series
# The R code need to use library stats and qcc
# The data ldeaths been used for this application is
# in Library stats
require(stats)

# library is used to generate the Control Charts
require(qcc)

#=========================================================
# preliminary analysis for the ldeaths data
# plot the original death data separate by male,
# female, and the total
# Figure 4.4: Death data: Monthly deaths from lung diseases
# in the UK

ts.plot(ldeaths, mdeaths, fdeaths, lty = c(1, 3, 4),
    xlab = "year", ylab = "deaths")
msg <- c('Total','Male','Female')
legend(1978,3900,lty=c(1,3,4), msg, bty='n')

# Autocorrelation plots for the multiple time series
# of male and female deaths
# Figure 4.5:  Autocorrelation and autocovariance plots
```

```
# for the Deaths data
par(mfrow=c(1,2))
acf(ldeaths, main='autocorrelation')
acf(ldeaths, type = 'covariance', main='autocovariance')

# Figure 4.6:  Spectral density estimates for the Deaths data
par(mfrow = c(2, 2))
spectrum(ldeaths, main='Series:deaths\n Raw Periodogram')
spectrum(ldeaths, spans = c(3, 3),
    main='Series:deaths\n Smoother = 5')
spectrum(ldeaths, spans = c(5, 7),
    main='Series:deaths\n Smoother = 7')
cpgram(ldeaths, main='Series:deaths')

#============================================================
# decompose the Total data into seasonal, trend,
# and remainder three parts
deaths <- stl(ldeaths, "periodic")
seasonal <- deaths$time.series[,1]
trend <- deaths$time.series[,2]
remainder <- deaths$time.series[,3]

# Figure 4.7:  The decomposition for the Deaths data
ts.plot(remainder+mean(ldeaths), seasonal, trend, ldeaths,
        lty=c(1,2,3,4), xlab='year', ylab='deaths')
msg <- c('Non-seasonal data','Seasonal component',
    'Trend','Original data')
legend(1977,4000,lty=c(1,2,3,4), msg, bty='n')


#============================================================
# Step 2: Compute the standard ewma without risk adjustment
# without seasonal effect
# The standard EWMA smooth with decay parameter kappa
# The decay parameter is set as .925
kappa <- 0.925

x <- time(ldeaths)
y <- trend + remainder
n <- length(y)

# std.ewma: the standard ewma value of ldeaths
# This standard ewma with seasonal effect
std.ewma <- ewmaSmooth(x,y,lambda=(1 - kappa),start=mean(y))
#lines(std.ewma, col="red")

#============================================================
# Compare deseasonal data (with trend),  Compare among:
# (a) The deseasonal data
# (b) Trend Component
```

```
# (c) Standard EWMA by using deseasonal data

# Figure 4.8:  The non-seasonal component, trend,
# and standard EWMA for the Deaths data
ts.plot(remainder+trend, trend, lty=c(2,3),
    xlab='year', ylab='deaths')
lines(std.ewma, lty=1)
msg <- c('Non seasonal data','Trend component',
    'EWMA without seasonal effect')
legend(1976.5,3000,lty=c(2,3,1), msg, bty='n')

#=========================================================
# Apply the deseasonal data (y =  trend + remainder)
#to EWMA control chart
x <- time(ldeaths)
y <- remainder  + trend
n <- length(y)
sample <- 1:n

# Figure 4.9:  the EWMA control chart for the Remainder
# with trend

q <- qcc(remainder  + trend, type="xbar.one", plot=FALSE)
qcc.options(bg.margin="white")
ewma(q,xlab='time',ylab='deaths')
points(sample,trend,type='l',lty=3)
msg <- c('EWMA without seasonal effect','Trend component')
legend(30,3000,lty=c(1,3), msg, bty='n')

#=========================================================
# Compare the remainder component of the deaths data
# (without trend, without seasonality),
#
# Compare among:
# (a) The remainder
# (b) Standard EWMA by using the remainder
# (c) Risk Adjusted EWMA (RA_EWMA) by using

#   ==>> ( Remainder) as the data
#   ==>> The transformation of seasonal component
# as the risk factor

# (b) Standard EWMA by using the original data
kappa <- 0.925
x <- time(ldeaths)
y <- remainder

# std.ewma: the standard ewma value of ldeaths
# This standard ewma with seasonal effect
std.ewma <- ewmaSmooth(x,y,lambda=(1 - kappa),start=mean(y))
```

```
std.ewma <- std.ewma$y +mean(ldeaths)

#============================================================
# (c) Risk Adjusted EWMA (RA_EWMA)

mu0 <- mean(y)
n <- length(y)

phi <- 1
# a simple transformation for the seasonal component
# which used as risk-adjustment parameter.
delta <-log( round(seasonal/100 + 8))

# setup the initial values for code
muhat <- muhatplus <- NULL
ytilde <- NULL
r <- NULL
r0 <- mu0
s <- (phi)^(-1) * (1-kappa)^(-1)

# initialize the values and parameters at t = 0
t <- 0
muhat[t+1] <- r0/s
muhatplus[t+1] <- muhat[t+1]  * exp(delta[t+1])

ytilde[t+1] <- y[t+1] - (muhatplus[t+1] - muhat[t+1]  )
r[t+1] <- kappa * r0 + phi^(-1) *  ytilde[t+1]

#calculate the estimated values at t  from 1 to (n-1)
for (t in (1:(n-1))){
    muhat[t+1] <- r[t]/s
    muhatplus[t+1] <- muhat[t+1]  * exp(delta[t+1])
    ytilde[t+1] <- y[t+1] - (muhatplus[t+1] - muhat[t+1]  )
    r[t+1] <- kappa * r0 + phi^(-1) *  ytilde[t+1]
}

ra.ewma <- muhat+ mean(ldeaths)

# Figure 4.10:  the non-seasonal EWMA control chart
# without trend
remainders <- y
q <- qcc(remainders, type="xbar.one", plot=FALSE)
qcc.options(bg.margin="white")
ewma(q,xlab='time',ylab='deaths',lambda=(1-kappa))
points(ra.ewma-mean(ra.ewma),type='l',lty=2)
msg <- c('The remainder','RA_EWMA','Standard EWMA' )
legend(40,800,lty=c(3,2,1), lwd = c(3,1,1),msg, bty='n')

# Figure 4.11:  the EWMA and RA-EWMA for remainder component
plot( remainder+ mean(ldeaths), type='p', pch=20, xlab='year',
```

```
     ylab='deaths', ylim=c(1950,2200))
points(x,ra.ewma,type='l', lty=1)
points(x,std.ewma,type='l', lty=2)

msg <- c('The remainder','RA_EWMA','Standard EWMA' )
legend(1976.5,2200,lty=c(3,1,2), lwd = c(3,1,1),msg, bty='n')
```

Bibliography

Bloomfield, P. (2000), *Fourier Analysis of Time Series: An Introduction* (2nd ed.), New York: John Wiley and Sons.

Brockwell, P., and Davis, R. (2002), *Introduction to Time Series and Forecasting* (2nd ed.), New York: Springer-Verlag.

Carey, R., and Lloyd, R. (2001), *Measuring Quality Improvement in Healthcare* (1st ed.), Milwaukee: ASQ.

Cleveland, R., Cleveland, W., McRae, J., and Irma T. (1990), "A Seasonal-Trend Decomposition Procedure Based on Loess", *Journal of Official Statistics*, 6, 3-73.

Crowder, S. (1987), "A simple Method for Studying Run-Length Distributions of Exponentially Weighted Moving Average Charts", *Technometrics*, 29, 401-407.

Diggle, P. (1990), *Time Series: A Biostatistical Introduction* (1st ed.), Oxford: Science Publications.

Grigg, O., and Spiegelhalter, D. (2007), "A Simple Risk-Adjusted Exponentially Weighted Moving Average", *Journal of the American Statistical Association*, 102 (477).

Harvey, A. (1991), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, UK: Cambridge University Press.

Harvey, A., and Fernandes, C. (1989), "Time Series Models for Count or Qualitative Observations", *Journal of Business and Economic Statistics*, 7, 407-417.

Lucas, J., and Saccucci, M. (1990), "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements", *Technometrics*, 32, 1-29.

McCullagh, P., and Nelder, J. (1983), *Generalized Linear Models* (2nd ed.), London: Chapman & Hill.

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

Parsonnet, V. *et al.* (1989) "A Method of Uniform Stratification of Risk for Evaluating the Results of Surgery in Acquired Adult Heart Disease, *Circulation,* 79, 3-12.

Roberts, S.W. (1959), "Control Chart Tests Based on Geometric Moving Averages", *Technometrics*, 1, 239-250.

Steiner, S., Cook, R., Farewell, V., and Treasure, T. (2000), "Monitoring Surgical Performance Using Risk-Adjusted Cumulative Sum Charts", *Biostatistics*, 1, 441-452.

West, M., and Harrison, J. (1997), *Bayesian Forecasting and Dynamic Models* (2nd ed.), New York: Springer-Verlag.

Venables, W. N., and Ripley, B. D. (2002), *Modern Applied Statistics with S* (4th ed.), New York: Springer-Verlag.