CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCIES

MSc THESIS

Ulduz MAMMADOVA

GENERALIZED LINEAR MODEL BASED CONTROL CHARTS WITH POISSON RESPONSE

DEPARTMENT OF STATISTICS

ADANA, 2017

ABSTRACT

MSC THESIS

GENERALIZED LINEAR MODEL BASED CONTROL CHARTS WITH POISSON RESPONSE

Ulduz MAMMADOVA

CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCIES
DEPARTMENT OF STATISTICS

Supervisor : Prof. Dr. Mahmude Revan ÖZKALE
Year : 2017, Pages : 94
Jury : Prof. Dr. Mahmude Revan ÖZKALE
: Prof. Dr. Selahattin KAÇIRANLAR
: Doç. Dr. Tolunay GÖÇKEN

The control chart is an important tool in statistical process control. Detecting the false alarm quickly, the control chart has a good performance and it is important in industrial processes. Generalized Linear Model based control charts use response variables that follow the exponential distributions.

Performance of control charts are obtained by using MATLAB R2015a. The deviance residuals obtained from Poisson regression model are used as a control chart statistics. It is determined through simulation results that the performance of CUSUM control chart is the best. Shewhart control chart shows a good performance when the mean is shifted. EWMA control chart has worst performance in both cases.The plastic plywood data used in real case study. The CUSUM control chart quickly detects out-of-control observations.

Key words: Generalized Linear Models, Poisson regression, Control chart, ARL

ÖZ

YÜKSEK LİSANS TEZİ

POİSSON YANIT DEĞİŞKENLİ GENELLEŞDİRİLMİŞ LİNEER MODELLERE DAYALI KONTROL GAFİKLERİ

Ulduz MAMMADOVA

ÇUKUROVA UNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI

Danışman : Prof. Dr. Mahmude Revan ÖZKALE
Yıl: 2017, Sayfa: 94
Jüri : Prof. Dr. Mahmude Revan ÖZKALE
: Prof. Dr. Selahattin KAÇIRANLAR
: Doç. Dr. Tolunay GÖÇKEN

Kontrol grafikleri, İstatistiksel Süreç Kontrolündeki önemli araçlardan biridir. Yanlış alarmı hızlıca tespit eden kontrol grafiği iyi bir performansa sahiptir ve endüstriyel proseslerde önemlidir. Genelleştirilmiş Lineer Modellere dayalı kontrol grafikleri yanıt değişkenleri üstel dağılıma sahipken kullanır.

Kontrol tablolarının performansını elde etmek için MATLAB R2015a kullanılmıştır. Poisson regresyon modelinden elde edilen sapma kalıntıları bir kontrol grafiği istatistiği olarak alınmıştır. Simülasyonların sonucunda, regresyon katsayılarının kayması durumunda, CUSUM kontrol grafiğinin performansının en iyi olduğu saptanmıştır. Shewhart kontrol grafiğinin, ortalama kayma yapıldığında iyi bir performans gösterdiği belirlenmiştir. Diğer kontrol grafiklerinden farklı olarak EWMA kontrol grafiğinin her iki durumda da en kötü performansa sahip olduğu anlaşılmıştır. Plastik kontrplak verisi uygulamada kullanılmış, CUSUM kontrol grafiği kontrol dışı gözlemi daha erken tespit etmiştir.

Anahtar kelimeler: Genelleştirilmiş lineer modeller, Poisson regresyon, Kontrol grafiği, Ortalama koşum uzunluğu

EXTENDED ABSTRACT

Statistical process control (SPC) has been used to great effect in the manufacturing industry to increase productivity in processes by specifically identifying and reducing variation. A control chart is one of the basic and the most reliable tool of statistical quality control and is a graphical appearance of the process widely used to study the way a process changes over time. Control charts indicate upper and lower control limits and often include a center line. If all data points are within the control limits, variations in the values may be due to a common cause and process is said to be in-control. If data points fall outside the control limits, variations may be due to a special cause and the process is said to be out-of-control.

There are several types of control charts. Shewhart, CUSUM and EWMA control charts are mostly used control charts. CUSUM and EWMA control charts are much more sensitive to changes than Shewhart control charts.

Mostly the performance analysis of the control charts is examined by their average run length (ARL) value. A number of the control chart statistics until the first one which exceeds the control limit is the run length. The average of the run lengths is ARL.

The regression based control charts are the combinations of control chart techniques and regression models.The control charts based on regression model constructs by using identically and independently normally distributed residuals obtained from data.

Generalized Linear Models (GLMs) are a generalization of the ordinary linear regression which is one of the most often used tools in predictive modeling. It allows people to analyze the relationships between dependent and independent parameters. GLMs allow the response variable to come from a distribution other than the normal

distribution such as Poisson, binomial, gamma distributions. When the response variables follow the Poisson distribution GLM based control chart could be used instead of the Shewhart attribute charts for count data. Poisson Shewhart, Poisson CUSUM, Poisson EWMA control charts are basic methods to monitor the Poisson distributed observations. For monitoring process, deviance residuals are obtained from the regression model and used as a control chart statistics. The deviance residuals are commonly used residual type in this process since deviance residuals are asymptotically normally distributed.

MATLAB R2015a was the tool that was used for the simulations and creating control charts for real data. Simulations proceed under several types of shifts. ARL is used for analyzing the performance of control charts. As a result, every single control chart has good performances for some situations and bad performances for some.

The overall result indicates that the Poisson CUSUM control chart is at good accuracy in detecting changes in count response. However, in general, the CUSUM control chart has thebest performance in most of the shift combinations.

The data for real case study is consisted of evaluating the effect of input variables over the number of defects found in produced plywood. The response variable is the number of defects per laminated plastic plywood area and the predicted variables is volumetric shrinkage, assembly time, wood density, and drying temperature.

The Poisson Shewhart control chart, Poisson CUSUM control chart, and Poisson EWMA control chart are constructed by using the deviance residuals obtained from data.

The CUSUM control chart quickly detects the out-of-control observation. The Shewhart control chart is also detected the out-of-control observations whereas, the

EWMA control chart shows that the control chart statistics are not exceeded control limits, in other words, all observations are in control.

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Dr. Mahmude Revan ÖZKALE for her excellent guidance, support and encouragement throughout preparation of this thesis. I sincerely thank all staff of Department of Statistics at Çukurova University. A special thanks to my parents for supporting me spiritually throughout my education and my life in general. I would also like to thank my friends, especially to Özgür ÖZEL. I wouldn't have succeeded my thesis without their support and encouragement.

CONTENT **PAGE**

ABSTRACT..... I
ÖZ..... II
EXTENTED ABSTRACT..... III
ACKNOWLEDGEMENT..... VII
CONTENTS..... IX
LIST OF TABLES..... XIII
LIST OF FIGURES..... XV
LIST OF ABBREVIATIONS..... XVII
1. INTRODUCTION……..... 1
1.1. Literature Review..... 3
2. GENERALIZED LINEAR MODELS..... 9
2.1. Elements of Generalized Linear Models..... 9
2.2. Parameter Estimation..... 13
2.2.1. Maximum Likelihood Method..... 13
2.2.2. Iterative Methods..... 15
2.3. Hypothesis Test..... 18
2.3.1 Likelihood Ratio Test..... 18
2.3.2. Wald Test..... 19
2.3.3. Score Test..... 19
2.4. Goodness of Fit……..... 20
2.4.1. Deviance Statistic..... 21
2.4.2. Pearson Chi-Square Test..... 22
2.4.3. Pseudo R Square……..... 22
2.5. Residual Analysis in GLM……..... 22

3. POISSON REGRESSION MODELS..... 25
3.1. Poisson Distribution..... 25
3.2. Model For Poisson Regression…..... 26
3.3. Parameter Estimation of Poisson Regression Model..... 28
3.3.1. Estimation of Parameters..... 28
3.3.2. Testing Estimated Parameters..... 29
3.4. Statistical Inference for Poisson Regression Model..... 30
3.4.1. Goodness of Fit…….... 30
3.4.2. Residual Analysis..... 31
4. CONTROL CHARTS..... 33
4.1. Control Chart..... 33
4.1.1. Shewhart Control Chart for Variables..... 35
4.1.2. Shewhart Control Chart for Attributes..... 40
4.1.2.1. C Chart..... 40
4.1.2.2. U Chart..... 40
4.1.2.3. Np Chart..... 41
4.1.2.4. P Chart..... 41
4.1.3. CUSUM Control Chart..... 42
4.1.3.1. Tabular CUSUM..... 43
4.1.3.2. V-Mask CUSUM..... 44
4.1.4. EWMA Control Chart..... 46
4.2. Generalized Linear Model Based Control Charts..... 48
4.2.1. GLM Control Charts……..... 48
4.2.2. GLM Control Charts with Poisson Response..... 50
4.2.2.1. Poisson Shewhart Control Chart..... 51
4.2.2.2. Poisson CUSUM Control Chart..... 51

4.2.2.3. Poisson EWMA Control Chart..... 52
4.3. Measuring Performance of Control Chart..... 53
4.3.1. Average Run Length..... 54
4.3.1.1. ARL for Shewhart Control Chart..... 55
4.3.1.2. ARL for CUSUM and EWMA Control Charts..... 55
5. APPLICATION ILLUSTRATION..... 57
5.1. Empirical Study..... 57
5.1.1. Performance Analysis of Control Charts..... 59
5.1.1.1. Changes in the Regression Coefficients..... 60
5.1.1.2. Changes in the Mean…..... 65
5.2. Real Case Study..... 66
6. CONCLUSIONS..... 71
REFERENCES..... 73
CURRICULUM VITAE..... 81
APPENDICES..... 83

LIST OF TABLES **PAGE**

Table 2.1. The link functions for some distributions of exponential family (McCullah and Nelder, 1989)……..……..... 13
Table 5.1. The *ARL* performance of control charts under shift in $\beta_0 \pm \delta\sigma$..... 60
Table 5.2. The *ARL* performance of control charts under shift in $\beta_1 \pm \delta\sigma$..... 60
Table 5.3. The *ARL* performance of control charts under shift in $\beta_2 \pm \delta\sigma$..... 61
Table 5.4. The ARL performance of control charts under shift in $\beta_0 \pm \delta\sigma$ and $\beta_1 \pm \delta\sigma$..... 62
Table 5.5. The ARL performance of control charts under shift in $\beta_0 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$..... 62
Table 5.6. The ARL performance of control charts under shift in $\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$..... 63
Table 5.7. The *ARL* performance of Shewhart control chart under shift in $\beta_0 \pm \delta\sigma$,$\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$..... 64
Table 5.8. The *ARL* performance of CUSUM control chart under shift in $\beta_0 \pm \delta\sigma$,$\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$..... 64
Table 5.9. The *ARL* performance of EWMA control chart under shift in $\beta_0 \pm \delta\sigma$,$\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$..... 65
Table 5.10. The *ARL* performance of control charts under shift in $\lambda \pm \delta\sigma$..... 65
Table 5.11. Deviance residuals obtained from data..... 68

LIST OF FIGURES **PAGE**

Figure 1.1. Representation of categorized papers using model based control charts (Asadzadeh et al., 2008) 4
Figure 4.1. Typical V mask CUSUM (Montgomery, 2009)..... 45
Figure 5.1. Kolmogorov-Smirnov test result of response variables..... 67
Figure 5.2. Shewhart control chart..... 69
Figure 5.3. CUSUM control chart..... 69
Figure 5.4. EWMA control chart..... 70

LIST OF ABBREVIATONS

ARL	: Average Run Length
RL	: Run Length
CUSUM	: Cumulative Sum
EWMA	: Exponentially Weighted Moving Average
GLM	: Generalized Liner Model
ML	: Maximum Likelihood
MLE	: Maximum Likelihood Estimation
CSC	: Cause Selecting Chart
SPC	: Statistical Process Control
SDRL	: Standard Deviation of Run Length
ANOS	: Average Number of Observations
ATS	: Average Time to Signal

1. INTRODUCTION

Statistical Process Control (SPC) is a set of statistical methods that are widely used to monitor, control and improve processes. One of the primary techniques of SPC is the control chart, which is employed to study the process variation that may affect the quality of the process outcome. Control charts were first introduced by Shewhart at Bell Labs, formerly known as Western Electric, in the early 1920s. The goal of control charts is to automatically determine if a deviation in a process is due to common causes, e.g., input fluctuation, or due to special causes, e.g., defects. Control charts were originally used to monitor deviation on telephone switches. Control charts have since become a common tool to detect problems in manufacturing processes where raw materials are inputs and the completed products.

After years of development and improvement, the practical applications of control charts extendto many areas such as engineering design, environmental science, pharmaceutical process, finance, accounting, and marketing (see, e.g., Ryan, 2000; Montgomery, 2012; Stoumbos et al., 2003).

A regression based control chart is the combination of the control chart techniques and simple linear regression models and was firstly proposed by Mandel (1969). Hawkins (1991, 1993) proposes monitoring the residuals from the process variable of interest when the variable is regressed on all the others. Residuals should be identically and independently normally distributed with zero mean and unit variance if process parameters are well estimated. The main advantage of a residual chart is that it can be applied to any autocorrelated data whether the process is stationary or not. Generalized Linear Model (GLM) based control chart

is same with the regression based control charts and applies when response data are exponentially distributed.

Average and standard deviation of run length, the average time for the signal are methods for measuring the performance of control charts. Average run length (ARL) is the average number of points that must be plotted before a point indicates an out-of-control condition and the most common used metric for evaluating the performance of a control chart.

The main objective of the thesis is to determine which type of control charts gives the best overall statistical performance when response variables follow the Poisson distribution. The control charts considered in this thesis are Shewhart, Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA) control charts. The control chart statistics are calculated from the deviance residuals which are obtained from Poisson regression model. Upper and lower control limits for control charts are calculated using the residuals. ARL is used for comparison of control charts.

The thesis is composed of several parts and organized as follows.

Chapter 2 covers the general information about GLM. Parameter estimation methods as maximum likelihood (ML) and iterative methods, calculation of test statistics - Likelihood ratio, Wald, and Score test statistics for testing hypotheses, goodness of fit tests, and residual analysis are given in this chapter.

In Chapter 3, Poisson Regression Model, which is one of the GLMs, is described. Poisson distributions, parameter estimation by using ML and iterative methods, the goodness of fit of the model and residual analysis are discussed.

Chapter 4 describes control charts and extension of control chart to GLM. This chapter covers the types of control charts, control charts based on GLM

especially with Poisson distributed response and method for a measure of performance of control charts.

Simulation study and results and a real case studyare given in Chapter 5. Major findings are discussed with regard to the research objectives.

1.1. Literature Review

The earliest reference to regression control charts in the quality control literature is appeared in DiPaola (1945). In his work, the control chart was described with regard to the simultaneous control of two variables.This type of control problem was later treated more extensively (Jackson, 1956; Weis, 1957). However, many authors claim that the model based control strategy was originated in Mandel (1969).

Mandel (1969) discussed in detail the combination of the conventional control chart and regression analysis in order to control a varying rather than a constant mean and developed a regression control chart to work in conjunction with administrative applications. Mandel (1969) designed a regression control chart by fitting regression lines to historical data and establishing limits around the regression line. He also shows the difference between a conventional control chart and a regression control chart in several respects. The usefulness of control chart is indicated by the variety of applications in which it can be used.

Zhang (1984) proposed the cause selected charts (CSCs) which are, in fact, similar to the regression control chart of Mandel (1969) in terms of charting a variable after the observations have been adjusted for the effect of an outside covariate. The CSC was implemented in the automobile industry which revealed the high efficiency of this type of control chart in comparison with the common Shewhart control charts for a given output quality characteristic.

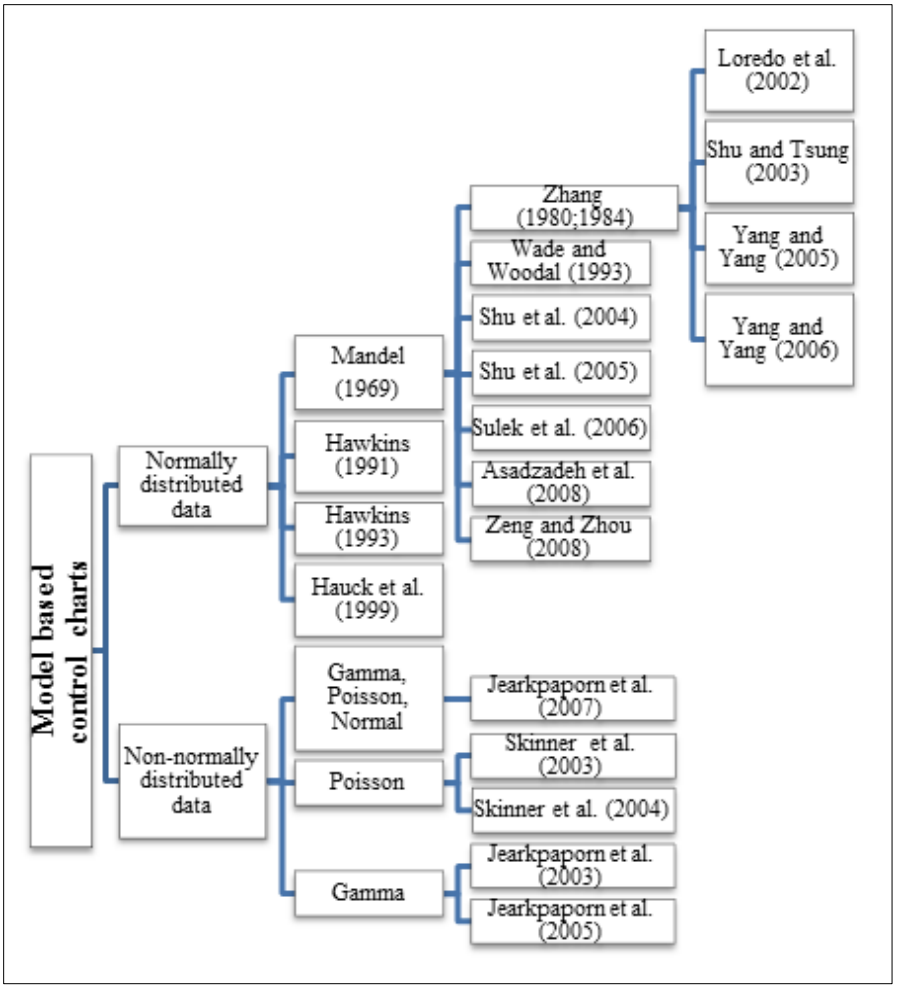


Figure1.1. Representation of categorized papers using model based control charts (Asadzadeh et al., 2008)

Wade and Woodall (1993) reviewed the basic principles of the CSC in the simple case of a two-step process and gave an example to illustrate its use. Shu et al. (2004) investigated the effect of parameter estimation errors on the performance of CSCs. After parameter estimation, the findings indicate that the charted statistics are correlated. Shu

et. al. (2004) discussed the run length performance of regression control charts based on estimated parameters for the manufacturing process. May et al. (2008) proposed an alternative regression control chart based on least absolute value regression for limited process data and provided a series of control charting procedures used to identify the variables that are significant to an out-of-control condition.

Hawkins (1991) introduced two types of regression adjustment: the model-void and the model fixed or cascade procedure. The model-void procedure involves the situation where some or all of the quality characteristics are related, the shift in one of the characteristics doesn't affect the others. The cascade procedure involves the process that has a natural ordering in which any characteristics shift may affect some or all of the measures following it. For the model-void process, Hawkins (1993) suggested regressing x_j on all the variablesx_i, $j \neq i$ where *x* donates the quality characteristic of interest. Residuals are computed from the model and used in monitoring the process. Model-fixed or cascade process – shift in any quality characteristic may affect some or all of the measures. The residuals are obtained from regressing x_j on$x_1, x_2, \dots, x_{j-1}$. The residuals will follow the normal distribution when the process is in control in both procedures. In his works, Hawkins (1991, 1993) also discussed advantages and disadvantages of traditional multivariate control charts. If the proper set of the variables is included to the model, the residuals from the model are usually uncorrelated over the time, even though the original variable of interest shows auto-correlation. This is the attractive feature of regression adjustments control chart.

Loredo (2002) study a method for monitoring autocorrelated processes based on regression adjustment. It is very useful to use residual based control chart when the input variables are observable and the process has a relationship between input and output variables. The performance of the residual based control chart in terms of the ARL is

compared to observation based control charts via Monte Carlo simulations. As a result, the observation based control charts perform very poorly when data are correlated over time. Under the assumption that the model is correct, the residual based control charts are superior for all considered cases. The advantages of residual based regression control charts are the ease of interpretation and straight forward implementation that requires only a least squares regression. An additional advantage is that the regression based control chart has a flexibility in modeling the relationship between the input and output variables. For small shifts, the CUSUM charts detect the shift more quickly than EWMA and Shewhart charts. The proposed method is recommended particularly for chemical processes where there are often cascade processes with several inputs but only a few outputs, and where many of the variables are highly autocorrelated.

Although all the previous methods are effective, those papers are based on the assumption that the monitored quality characteristics are normally distributed. But generally, quality characteristics do not always follow this assumption. In this case, a modification of the model based control strategy using a generalized linear modeling technique is proposed.

GLMs extend the regression framework to situations where the distribution of the response is the member of the exponential family (Asadzadeh et al., 2008).

Skinner (2003) proposed a procedure which is based on the likelihood ratio statistic, i.e., the deviance residual resulting from a GLM for Poisson counts. The deviance residual based control chart is compared with univariate and bivariate *C* charts and the deviance residual based control chart shows an effective result than a *C* charts. Skinner (2004) also compared the deviance residual based control chart with *C* chart for overdispersed Poisson data. *C* chart outperforms the GLM based control chart since GLM used to calculate the deviance residual effectively removes the impact of the input

variables, thus eliminates the overdispersion from the control statistic when overdispersion interferes with the *C* chart.

Jearkpaporn (2003) used a model based control chart for monitoring gamma distributed multiple variables. Under the assumption that the shape parameter is unchanged and the input and output variables are related in a certain manner like the deviance residual. As a result of comparison of an individual Shewhart chart and the deviance residual based control chart, the last one detects the shifts more quickly. The deviance residual based control chart also shows a good ARL in small shifts.

Amiri et al. (2014) propose a method for monitoring the two stage processes with Poisson variable. For process monitoring, they used a new link function which proposed to establish the relationship between response and input variables using a regression function. Proposed by the author new link function is a combination of two existing link functions - square root and log link functions which show the relationship between Poisson response and predictor variables, assuming that the parameters are known. The statistic is obtained by using the standardized residuals and then they used Shewhart and EWMA control charts to monitor the proposed statistic. The ARL is used for the comparison of the proposed method to the method proposed by Skinner (2003). The comparisons were made based on three types of shifts. The results show that the proposed method shows a better performance.

Filho (2015) used control charts for industrial processes that involve the count variable. The control limits for constructing of the control chart obtained by using the deviance residuals from the regression model. The ARL of control chart calculated for in-control and out-of-control scenarios. The overall results show the residual based control chart has a good accuracy in detection changes in count response. His work also represents the combination of the Poisson regression and principal component analysis

for monitoring the manufacturing process. The process is evaluated by using the control chart of deviance residuals from the model. The new approach shows a good performance.

2. GENERALIZED LINEAR MODELS

The ordinary linear regression models used for modeling the relationship between a scalar response variable *y* and one or more predictor variables *x* usually assumes that the response variable has a normal distribution.

GLM was developed by Nelder and Wedderburn (1972). These models are used for modeling the relationship between the predictor and response variables when the response variables are followed distributions from the exponential family. The exponential family covers a large number of distributions. Discrete distributions as the Poisson, binomial, and Bernoulli or continuous distributions as the normal, Gamma or Inverse Gaussian distribution are members of this family.

2.1. Elements of Generalized Linear Models

The GLM consists of three elements.

Random component of a GLM defines the probability distribution of the response variable. The probability density functions of distributions which belong to the exponential family are usually described as*:*

$$f(y_i;\theta_i,\varphi) = \exp\left(\frac{y_i\theta_i - b(\theta_i)}{a(\varphi)} + c(y_i,\varphi)\right) \qquad (2.1.)$$

where *y* is a response variable, $a(\varphi)$, $b(\theta_i)$ and $c(y_i,\varphi)$ are the unknown functions.

The function $a(\varphi)$ is commonly of the form

$$a(\varphi) = \varphi/w,$$

where φ, usually donated by σ^2 and called the dispersion parameter, and constant over observations and w is a known prior weight that varies from observation to

observation. The parameter θ_i is a scale parameter. θ_i and φ are orthogonal parameters for exponential distributions.

The mean and the variance of distributions in exponential family is

$$E(y_i) = \mu_i = b'(\theta_i)$$
$$var(y_i) = \sigma_i^2 = b''(\theta_i)a(\varphi)$$

where $b'(\theta_i)$ and $b''(\theta_i)$ are the first and second order derivatives of $b(\theta_i)$ evaluated at θ_i.

The variance of y_i is the product of two functions; the $b''(\theta_i)$, depends on the canonical parameter (and hence on the mean) only and will be called variance function while other is independent of θ_i and depends only on φ. The variance function of μ will be written *V(μ)*. In other words, a variance function describes how the variance, $var(y_i)$ depends on the mean. When $a(\varphi) = \varphi/w_i$ the variance is

$$var(y_i) = \sigma_i^2 = \varphi b''(\theta_i)/w_i$$

The mean and the variance could be obtained using the well-known relations

$$E\left(\frac{\partial L_i}{\partial \theta_i}\right) = 0 \qquad (2.2.)$$

and

$$E\left(\frac{\partial^2 L_i}{\partial {\theta_i}^2}\right) + E\left(\frac{\partial L_i}{\partial \theta_i}\right)^2 = 0 \qquad (2.3.)$$

where $L_i = \log f(y_i; \theta_i, \varphi)$.

The differentiation of L_i with respect to θ_i is

$$\frac{\partial L_i}{\partial \theta_i} = \frac{y_i - b'(\theta_i)}{a(\varphi)}$$

and

$$\frac{\partial^2 L_i}{\partial {\theta_i}^2} = -\frac{b''(\theta_i)}{a(\varphi)} \tag{2.4.}$$

Using Eq. (2.2.) we get

$$E\left(\frac{\partial L_i}{\partial \theta_i}\right) = E\left(\frac{y_i - b'(\theta_i)}{a(\varphi)}\right) = 0$$

Since $a(\varphi)$ is a fixed constant, $E(y_i) - b'(\theta_i) = 0$, so

$$E(y_i) = b'(\theta_i) \tag{2.5.}$$

Similarly, using Eq. (2.3.), (2.4.) and (2.5.), we obtain

$$-\frac{b''(\theta_i)}{a(\varphi)} + \frac{var(y_i)}{a^2(\varphi)} = 0 \tag{2.6.}$$

so that

$$var(y_i) = b''(\theta_i)a(\varphi). \tag{2.7.}$$

The second component of the GLM is a *linear predictor*. Linear predictor describes how the location of the response distribution changes with the predictor variables.

Suppose that we have a vector of *q=p+1* unknown parameters $\beta = (\beta_0, \beta_1 \ldots \beta_q)'$ and a matrix of known predicted variables for the *n* observations *X*, the linear predictor is

$$\eta_i = x_i'\beta$$

The terms x_i and β describe the predictor variables and parameters, respectively. x_i is the $q \times 1$ vector of predictor variables

$$x_i' = \left[1\ x_{1i} \quad x_{2i} \ldots\ x_{pi}\right]$$

which is the *i*th row of the *X* matrix.

The *link function* is the third component of GLM and describes a relationship between the linear predictor and the mean of the distribution function. In other words, link function g(·) connects the random component and the linear predictor:

$$g(\mu_i) = \eta_i = x'_i\beta.$$

The function g(·) is monotone and differentiable function. The relationship between the mean and the linear predictor is

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(x'_i\beta). \quad (2.8.)$$

If the link function g(·) which transforms mean μ_i to the scale parameter θ_i, it is called as the canonical link function. Since, $\mu_i = g\big(b'(\theta_i)\big)$ the canonical link is defined as

$$g = \big(b'(\theta_i)\big)^{-1} \Longrightarrow g(\mu_i) = \theta_i$$

$$\theta_i = \sum_{j=0}^{q} \beta_j x_{ji}$$

and equates the scale parameter to the linear predictor. An advantage of canonical link function is that a minimal sufficient statistic for β exists, i.e. all the information about β is contained in a function of the data of the same dimensionality as β.

There are many commonly used link functions. The link functions can be chosen, depending on the distribution of the response variable.

Table 2.1. shows the scale parameters of some exponential distributions, commonly used link functions and models depend on the distribution of response variable.

Table 2.1. The link functions for some distributions of exponential family (McCullah and Nelder, 1989)

Distribution	**Scale parameter**	**Link Function**	
Normal	$\theta_i = \mu_i$	Identity link	$\mu_i = x'_i\beta$
Poisson	$\theta_i = \ln(\mu_i)$	Log link	$\ln(\mu_i) = x'_i\beta$
Binom	$\theta_i = \ln\left(\frac{\pi_i}{1-\pi_i}\right)$	Logit link	$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = x'_i\beta$

2.2. Parameter Estimation

The term parameter estimation refers to the process of using sample data to estimate the parameters of the selected distribution. Several parameter estimation methods are available.

In GLM, the parameter estimation is based on ML functions. The basic idea behind ML is to obtain the most likely values of the parameters, for a given distribution, that will best describe the data. The nonlinearity of likelihood equations makes hard estimating parameters. So, the estimators are obtained by using iterative methods like Newton-Raphson and Fisher's scoring.

2.2.1. Maximum Likelihood Estimation

Maximum likelihood estimation is one of the fundamental methods for constructing estimators of unknown parameters in statistical estimation theory. This method estimates the parameters of the statistical model by finding the parameter

values that maximize the likelihood of making the observations given the parameters. The ML estimation method also should be applied to the GLMs. Let $L_i = \log f(y_i; \theta_i, \varphi)$ and $l(\beta; y_i) = \sum_{i=1}^{n} L_i$. The log likelihood function for the sample $y_i, i = 1,2 \dots n$ is

$$l(\beta; y_i) = \sum_{i=1}^{n} L_i = \sum_{i=1}^{n} \log f(y_i; \theta_i, \varphi) + \sum_{i=1}^{n} c(y_i, \varphi). \qquad (2.11.)$$

First order derivative of log likelihood with respect to β_j is known as score function. The variance of the score is known as the Fisher information and is written *I*.

The likelihood equation for GLMs is

$$\frac{\partial l(\beta; y_i)}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial L_i}{\partial \beta_j} = 0 \text{ for all } j. \qquad (2.12.)$$

The differentiation of log likelihood using the chain rule is

$$\frac{\partial L_i}{\partial \beta_j} = \frac{\partial L_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j}.$$

The partial derivatives,

$$\frac{\partial L_i}{\partial \theta_i} = \frac{y_i - b'(\theta_i)}{a(\varphi)}$$

Since, $\mu_i = b'(\theta_i)$

$$\frac{\partial \theta_i}{\partial \mu_i} = \left(\frac{\partial \mu_i}{\partial \theta_i}\right)^{-1} = \frac{1}{b''(\theta_i)} = \frac{a(\varphi)}{var(y_i)}$$

$$\frac{\partial \eta_i}{\partial \beta_j} = \frac{\partial \sum_{j=0}^{q} \beta_j x_{ji}}{\partial \beta_j} = x_{ji}$$

Since $\eta_i = g(\mu_i)$ derivative of $\frac{\partial \mu_i}{\partial \eta_i}$ is depend on link function. Without determining link function this derivative cannot be defined.

So Eq. (2.12.) results in

$$\frac{\partial l(\beta; y_i)}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial L_i}{\partial \beta_j} = \sum_{i=1}^{n} \frac{y_i - \mu_i}{a(\varphi)} \frac{a(\varphi)}{var(y_i)} \frac{\partial \mu_i}{\partial \eta_i} x_{ji} = 0, j = 0, 1,2, \dots q. \quad (2.13.)$$

The likelihood equations are non-linear functions of β and must be solved by using the iterative methods.

2.2.2. Iterative Methods

In general, there are two iterative methods which are commonly used for estimating the parameters of a nonlinear equation.

1. Newton-Raphson method
2. Fisher's scoring method

Both methods have a same general form and differ only in the variance structure.

Newton – Raphson method is based on Taylor series expansion around some given point.

The nonlinear equation is solved by determining the point at which a function takes its maximum. It begins with an initial approximation for the solution. It obtains the second approximation by approximating the function in a neighborhood of the initial approximation by a second-degree polynomial and then finding the location of that polynomial's maximum value. This step repeats to generate a sequence of approximations. These converge to the location of the maximum when the function is suitable and/or the initial approximation is good (Sunashini, 2006).

The Newton-Raphson consists of the scheme

$$(\beta^{(t+1)})' = (\beta^{(t)})' - (H^{(t)})^{-1} u^{(t)} \quad (2.14.)$$

where *t* is a step in the iterative process *(t=0,1,2…),* the $q \times 1$ vector $u^{(t)}$ is

$$u^{(t)} = \left(\frac{\partial l(\beta)}{\partial \beta_0}, \frac{\partial l(\beta)}{\partial \beta_1}, \ldots, \frac{\partial l(\beta}{\partial \beta_q}\right)' \mid_{\beta=\beta^{(t)}}$$

where $l(\beta) = \sum_{i=1}^{n} L_i$ and the $q \times q$ Hessian matrix $H^{(t)}$ is given by

$$H_{jk}^{(t)} = \frac{\partial^2 l(\beta)}{\partial \beta_j \partial \beta_k} \mid_{\beta=\beta^{(t)}},$$

For *j, k=0, 1, 2, ...,q,* both $u^{(t)}$ and $H^{(t)}$ being evaluated at the current estimate $\beta^{(t)}$.

By using chain rule, the score function at the *i*th iteration is

$$u_j^{(t)} = \frac{\partial\, l(\beta)}{\partial \beta_j} \mid_{\beta=\beta^{(t)}} = \sum_{i=1}^{n} \frac{\partial L_i}{\partial \beta_j} \mid_{\beta=\beta^{(t)}}$$

$$= \sum_{i=1}^{n} \frac{dL_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta_j} \mid_{\beta=\beta^{(t)}} = \sum_{i=1}^{n} \frac{dL_i}{d\eta_i} \mid_{\beta=\beta^{(t)}} x_{ji}.$$

The Hessian at the *i*th iteration is

$$H_{jk}^{(t)} = \frac{\partial^2 l(\beta)}{\partial \beta_j \partial \beta_k} \mid_{\beta=\beta^{(t)}} = \sum_{i=1}^{n} \frac{\partial^2 L_i}{\partial \beta_j \partial \beta_k} \mid_{\beta=\beta^{(t)}}$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial \beta_k} \left(\frac{\partial L_i}{\partial \beta_j}\right) \mid_{\beta=\beta^{(t)}}$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial \eta_i} \left(\frac{\partial L_i}{\partial \eta_i} x_{ij}\right) \mid_{\beta=\beta^{(t)}}$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial \beta_k} \left(\frac{\partial L_i}{\partial \eta_i} x_{ji}\right) x_{ki}$$

$$= \sum_{i=1}^{n} \frac{\partial^2 L_i}{\partial {\eta_i}^2} \mid_{\beta=\beta^{(t)}} x_{ji} x_{ki}.$$

If write $u^{(t)}$ for the $n \times 1$ vector with

$$s_i^{(t)} = \frac{\partial L_i}{\partial \eta_i} \mid_{\beta=\beta^{(t)}}$$

and $W^{(t)}$ for the diagonal $n \times n$ matrix with

$$w_{ii} = -\frac{d^2 L_i}{d\eta_i{}^2}$$

we have $u^{(t)} = X's^{(t)}$ and $H = -X'W^{(t)}X$, than the Newton Raphson iteration results in

$$\begin{aligned}\left(\beta^{(t+1)}\right)' &= \left(\beta^{(t)}\right)' - \left(H^{(t)}\right)^{-1} u^{(t)} \\ &= \left(\beta^{(t)}\right)' + (X'W^{(t)}X)^{-1}X's^{(t)}\end{aligned}$$

In this equation $H^{(t)}$ has considered as nonsingular. Iterations proceed until changes in $l(\beta)$ between successive cycles are sufficiently small (Suhasini, 2006).

Fisher's scoring method same with least squares method in linear regression, but in this method reweighting is in every step. The method of scoring is a procedure that is very similar to Newton-Raphson. Fisher's scoring uses expected value of Hessian matrix which called the expected information matrix. The Hessian matrix itself called the observed information matrix.

Let I^t donate approximation t for the ML estimate of the expected information matrix that has elements $-\mathrm{E}(\partial^2 l(\beta)/\partial\beta_j \partial\beta_k)$, evaluated at $\beta^{(t)}$.

The Fisher's scoring iteration has a form

$$\beta^{(t+1)} = \beta^{(t)} - (I^t)^{-1} u^{(t)} \qquad (2.15.)$$

The iteration process for parameter estimation proceeds until convergence. The convergence is when difference $\beta^{(t+1)} - \beta^{(t)}$ smaller than k"convergence criterion", then $\beta^{(t+1)} = \hat{\beta}$. For GLM β is a vector, then the difference $\beta^{(t+1)} - \beta^{(t)}$ yields a vector. Convergence could be reached when all $|\beta^{(t+1)} - \beta^{(t)}| < k$ or when $\sum |\beta^{(t+1)} - \beta^{(t)}| < k$.

2.3. Hypothesis Test

The hypothesis tests are used to compare how well two related models fit the data. For GLMs, the two models should have the same probability distribution and the same link function but the linear predictor of one model has more parameters than the other. The simpler model, corresponding to the null hypothesis H_0, must be a special case of the more general model. If the simpler model fits the data as well as the more general model does, then it is preferred on the grounds of parsimony and H_0 is retained. If the more general model fits significantly better, then H_0 is rejected in favor of an alternative hypothesis H_1 which corresponds to the more general model (Dobson, 2008).

Hypothesis about β_j can be tested by using likelihood ratio, Wald or score test statistics. These test statistics are asymptotically equivalent test statistics.

For a generic scalar model parameter β the hypothesis is

$$H_0: \beta = \beta_0$$
$$H_1: \beta \neq \beta_0$$

where β_0 is the particular null value, typically 0.

2.3.1. Likelihood Ratio Test

Suppose that l_0 is the likelihood function calculated at the β_0, l_1 is the likelihood function calculated at the β which permitting H_0 or H_1 to be true. The likelihood ratio test is actually based on the likelihood ratio:

$$\Lambda = \frac{l_0}{l_1} = \frac{\prod_{i=1}^{n} f(y_i; \theta_i, \varphi)}{\max \prod_{i=1}^{n} f(y_i; \theta_i, \varphi)}.$$

The ratio $\Lambda = l_0/l_1 \leq 1$, since l_0 results from maximizing at a restricted β value.

The likelihood ratio test statistics is defined as follows:

$$-2\log\Lambda = -2\log\left(\frac{l_0}{l_1}\right) = -2(L_0 - L_1) \qquad (2.16.)$$

where L_0 and L_1 donate the maximized log-likelihood functions. It follows a limiting null chi-squared distribution as $n \to \infty$, with *df=1*. Large values of $log\Lambda$ suggest that the model of interest is a poor description of the data. The critical region for $log\Lambda$ could be determined by sampling distribution.

The likelihood ratio test rejects the null hypothesis if the value of this statistic is too small.

2.3.2. Wald Test

For $H_0: \beta = \beta_0$ the Wald statistics is

$$z = \frac{\hat{\beta} - \beta_0}{se(\beta)} \qquad (2.17.)$$

where $se(\beta)$ is the asymptotic standard error of the estimated coefficient β. Under the null hypothesis, *z* follows an approximately standard normal distribution and z^2 follows an approximate chi-squared distribution with *df=1*.

For multiple parameters $\beta = (\beta_0, \beta_1)$ to test the hypothesis $H_0: \beta_0 = 0$, the Wald test statistics is

$$z = \hat{\beta}_0' \, [\widehat{var}(\hat{\beta}_0)]^{-1} \hat{\beta}_0 \qquad (2.18.)$$

where $\hat{\beta}_0$ is an unrestricted maximum likelihood estimate of β_0 and $\widehat{var}(\hat{\beta}_0)$ is unrestricted estimated covariance matrix of $\hat{\beta}$.

2.3.3. Score Test

The score test is like the likelihood ratio test, is a general asymptotic parametric test associated with the likelihood function (Rao, 1947). Score tests are

often simpler than likelihood ratio tests because the statistic requires parameter estimators to be obtained only under the null hypothesis (Smyth, 2003).

The score test uses the slope and the expected curvature of the log-likelihood function, evaluated at the null value β_0. The chi-squared form of the score test is

$$\mathrm{S} = \frac{(\partial L(\beta)/\partial \beta_0)^2}{-\mathrm{E}(\partial^2 L(\beta)/\partial \beta_0^2)} \tag{2.19.}$$

where the notation reflects derivatives with respect to β that is evaluated at β_0. When hypothesis is true $s \sim \chi^2_{df=1}$.

For multiple parameters, a score statistic is a quadratic form based on the vector of partial derivatives of the log likelihood and inverse information matrix.

In other words, for *p* dimensional restricted parameter vector, β_0

$$s = u'(\hat{\beta}_0)[\mathrm{I}^{-1}(\hat{\beta}_0)]u(\hat{\beta}_0) \tag{2.20.}$$

When hypothesis is true $s \sim \chi^2_{df=p}$. It's the maximum likelihood estimate of the other parameters in β that are not restricted in the null hypothesis.

2.4. Goodness of Fit

In general, goodness of fit test refers to measuring how well do the observed data correspond to the fitted (assumed) model. The common goodness of fit measures for GLM are the Pearson and deviance statistics, which are weighted sums of residuals and can be calculated from the data and the fitted model. Both measures have approximate chi-square distributions under the hypothesis that the current model is appropriate for fixed number of combinations of predictor variables and large counts. This measure can be compared with the chi-squared distribution with *n* – *p* degrees of freedom, where *n* is a number of observations and *p* is the number of parameters that are estimated. The chi-squared distribution is likely to be a better

approximation for the sampling distribution of Pearson than for the sampling distribution of deviance. Also, pseudo R squared measures can be used.

2.4.1. Deviance Statistic

The deviance statistic is defined as twice the difference in log-likelihood between saturated and full model. Saturated model is the most complex model possible with the given data.

Let $L(\hat{\mu}; y)$ donates the maximum of the log likelihood for the model. Considered for all possible models the maximum achievable log likelihood is $L(y; y)$. This occurs for the most general model having the separate parameter for each observation and the perfect fit $\hat{\mu} = y$. This is the saturated model.

Let $\hat{\theta}_i$ be an estimator of θ_i for chosen model and $\tilde{\theta}_i$ be an estimator of θ_i for the saturated model. The maximized log likelihood functions for full and saturated model is

$$-2log\left[\frac{ML\ for\ full\ model}{ML\ for\ saturated\ model}\right] = -2[L(\hat{\mu}; y) - L(y; y)]$$

From Eq. (2.1.) we have

$$-2[L(\hat{\mu}; y) - L(y; y)] = 2\sum_i (y_i\,\tilde{\theta}_i - b(\tilde{\theta}_i))/a(\varphi) - 2\sum_i (y_i\,\hat{\theta}_i - b(\hat{\theta}_i))/a(\varphi)$$

If $a(\varphi) = \varphi/\omega_i$ then

$$2\sum_i \omega_i (y_i\,(\hat{\theta}_i - \tilde{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i))/\varphi = D(y; \hat{\mu})/\varphi = SD(y; \hat{\mu}) \qquad (2.21.)$$

where $SD(y; \hat{\mu})$ is the scaled deviance and $D(y; \hat{\mu})$ is the deviance statistic (Myers, 2010).

The null hypothesis is accepted in α significance level if $D(y; \hat{\mu}) > \chi^2_{\alpha,n}$. Large deviance value is the evidence of lack of fit.

2.4.2. Pearson Chi-Square Statistic

For a GLM with response y_i, mean μ_i and variance σ_i the Pearson chi-square statistic is

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\sigma}_i} \quad (2.22.)$$

where $\hat{\mu}_i$ and $\hat{\sigma}_i$ are estimates of $\mu_i$and σ_i. For a large n values $(n \to \infty)$, χ^2 is asymptotically ${\chi^2}_{n-p}$ distributed. Very small value of χ^2 suggest that the model fits the data too well.

2.4.3. Pseudo R square

Pseudo R^2 deviance based measure is used for model comparison and has the property that, when applied to the linear model, they match the interpretation of the linear model R-squared.

Assume that $D(y;\bar{y})$ is the deviance of intercept model, $D(y;\hat{\mu})$ is the deviance of fitted model and $D(\hat{\mu};\bar{y})$ is the explained deviance. The pseudo R^2 is defined as

$$R^2_{pseudo} = 1 - \frac{D(y;\hat{\mu})}{D(\hat{\mu};\bar{y})} \; or \; R^2_{pseudo} = \frac{D(\hat{\mu};\bar{y})}{D(y;\bar{y})} \quad (2.23.)$$

which goes from 0 to 1 where 1 shows the perfect fit.

2.5. Residual Analysis in GLM

The residual analysis is also important for GLMs. Analyzing the residuals produced by any statistical tool provides us with a general approach to assessing the quality and find possible problems with the data. The residuals should be plotted against each continuous predictor variable in the model to check if the assumption of linearity is appropriate and against other possible predictor variables not included in the model.

Several kinds of residuals can be defined for GLMs:

- The response residuals or raw residuals are differences between the observed response and its estimated expected value, $y_i - \hat{\mu}_i$ where $\widehat{\mu_i} = g^{-1}(\widehat{\eta_i}) = g^{-1}(\beta_0 + \beta_1 x_{1i} + \cdots + \beta_p x_{pi})$. The raw residuals are not the best choice for the GLM, because the approximate normality and constant variance assumptions are not satisfied.
- Pearson residuals depend on Pearson statistics. For a model with variance function $v(\mu)$, the Pearson residual for observation y_i and its fitted value $\widehat{\mu_i}$ is

$$e_i = \frac{y_i - \widehat{\mu}_i}{\sqrt{v(\widehat{\mu_i})}}. \tag{2.24.}$$

The sum of the squares of these residuals is the Pearson statistic. When $a(\varphi) = \varphi, \frac{e_i}{\sqrt{\varphi}}$ is called the *i*th scaled Pearson residual. In this case, the Pearson statistic is

$$\frac{1}{\varphi}\sum_{i=1}^{n} e_i = \frac{1}{\varphi}\sum_{i=1}^{n} \frac{(y_i - \widehat{\mu}_i)^2}{v(\widehat{\mu_i})}. \tag{2.25.}$$

- The deviance residuals are the square roots of the contribution to the deviance from each observation, multiplied by the sign of the corresponding raw residual

$$D_i = \sqrt{d_i} \times sign(y_i - \hat{\mu}) \tag{2.26.}$$

where $d_i = 2\omega_i[y_i(\widetilde{\theta_i} - \widehat{\theta_i}) - b(\widetilde{\theta_i}) + b(\widehat{\theta_i})]$.

A sum of squares of the deviance residuals is the deviance statistic.

- Standardized residual is defined as

$$r_i = \frac{y_i - \widehat{\mu_i}}{\sqrt{v(\widehat{\mu_i})(1 - \widehat{h}_{ii})}} = \frac{e_i}{\sqrt{1 - \widehat{h}_{ii}}} \tag{2.27.}$$

where $\hat{h}_{ii}$ is the estimation of h_{ii} which is the diagonal element of the generalized hat matrix H_w for observation *i.* The generalized hat matrix is

$$H_w = W^{1/2}X(X^TWX)^{-1}X^TW^{1/2}$$

where $W = diag(\partial\mu_i/\partial\eta_i)^2/var(y_i)$.

3. POISSON REGRESSION MODEL

The model for count data is the Poisson regression model. Poisson regression model is a nonlinear regression model.

3.1. Poisson Distribution

The Poisson regression model is a member of GLMs where the response variable follows the Poisson distribution.

Poisson distribution is a discrete distribution defined to express the number of events that occur in a unit of time or space. The possible values are nonnegative integers. In Poisson distribution, the trials must be very large while the probability of occurrences of outcome under observation must be small. In addition, the independence of trials and consistency of probability from trial to trial properties are required. The number of failures for a certain machine at various operating conditions, infant mortality or cancer incidence among groups with different demographics, the number of tropical cyclones during a season, the number of faulty components in a computer or in a batch of manufactured items are Poisson distribution examples.

The Poisson distributed data have to provide the following assumptions (Rodriguez, 2007):

- The probability of at least one occurrence of the event in a given time interval is proportional to the length of the interval.
- The probability of two or more occurrences of the event in a very small time interval is negligible.

- The numbers of occurrences of the event in disjoint time intervals are mutually independent.

The Poisson probability function for y_i ($i = 1,2, \ldots, n$) independent variables with μ_i mean is

$$f(y_i, \mu_i) = \frac{e^{-\mu_i} \mu_i{}^{y_i}}{y_i!}. \quad (3.1.)$$

The Poisson distribution has the property that the mean and the variance of distribution are equal

$$E(y_i) = var(y_i) = \mu_i \quad (3.2.)$$

3.2. Model for Poisson Regression

In GLMs, Poisson regression is used when the response variable follows the Poisson distribution. The Poisson regression models the relationship between the observed count and predictor variables. The model describes the mean of discrete distribution as a function of predicted variable.

A GLM for the set of independent variables y_i ($i = 1,2, \ldots, n$) which follows Poisson distribution with mean μ_i is

$$\begin{aligned} f(y_i, \mu_i) &= \frac{e^{-\mu_i} \mu_i{}^{y_i}}{y_i!} \\ &= exp[y_i log\mu_i - \mu_i - log(y_i!)] \\ &= exp[y_i\theta_i - exp(\theta_i) - log(y_i!)] \end{aligned} \quad (3.3.)$$

where scale parameter $\theta_i = log\mu_i$. $b(\theta_i) = \exp(\theta_i)$, $a(\varphi) = 1$ and $c(y_i, \varphi) = -log(y!)$ in Eq. (2.1.).

The mean and the variance of model are

$$E(y_i) = b'(\theta_i) = ex\,p(\theta_i) = \mu_i \quad (3.4.)$$

$$var(y_i) = b''(\theta_i)a(\varphi) = exp(\theta_i) = \mu_i \quad (3.5.)$$

The relationship between the mean of Poisson distributed data and the linear predictor can be shown by using the identity and log-linear models.

Identity link function is

$$g(\mu_i) = \mu_i = x'_i\beta. \quad (3.6.)$$

The relationship between the mean of the response variable and the linear predictor is

$$\mu_i = g^{-1}(x'_i\beta) = x'_i\beta.$$

But this model has the disadvantage that the linear predictor on the right-hand side can assume any real value, whereas the Poisson mean on the left-hand side, which represents an expected count, has to be non-negative (Rodriguez, 2007). The solution to this problem is to model instead of the logarithm of the mean using a linear model. By taking logarithm calculating $\eta_i = ln(\mu_i)$ and assume that the transformed mean have a linear model $\eta_i = x'_i\beta$. These two steps can be combined in one. We can write the log linear model as

$$ln(\mu_i) = x'_i\beta. \quad (3.7.)$$

In this model, β represents the expected changes in the log of the mean per unit change in the predictor *x*. In other words, increasing *x* is associated with increasing β in the log of the mean. To obtain the multiplicative model for the mean, we exponentiated Eq. (3.7.)

$$\mu_i = \exp(x'_i\beta).$$

This log link is more attractive for Poisson distribution because it ensures that all of the predicted values for response variable will be nonnegative.

3.3. Parameter Estimation of Poisson Regression Model

Basically, Poisson regression model is nonlinear regression model. But despite the nonlinearity, ML method is used to estimate parameters. Also, iterative methods are used for parameter estimation of the model.

3.3.1. Estimation Parameters

Assume that $(y_1, y_2, \dots, y_n)$ is an independent set of Poisson random variables with mean μ_i . Than the likellihood function in general form is

$$L(\beta; y_i) = \prod_{i=1}^{n} f_i(y_i, \mu_i) = \prod_{i=1}^{n} \frac{e^{-\mu_i} {\mu_i}^{y_i}}{y_i!} = \frac{\prod_{i=1}^{n} {\mu_i}^{y_i}\, e^{-\sum_{i=1}^{n} \mu_i}}{\prod_{i=1}^{n} y_i!} \qquad (3.8.)$$

where $\mu_i = \mathrm{g}^{-1}(x_i'\beta) = x_i'\beta$, *i=1,2,…n.*

Maximizing the likelihood function is equivalent to maximizing the log likelihood function

$$l(\beta; y_i) = \sum_{i=1}^{n} log\, f_i(y_i, \mu_i) = \sum_{i=1}^{n} y_i \log(\mu_i) - \sum_{i=1}^{n} \mu_i - \sum_{i=1}^{n} \log(y_i!)$$

Using the Eq. (3.6.) the log likelihood is

$$l(\beta; y_i) = \sum_{i=1}^{n} y_i {x_i}'\beta - \sum_{i=1}^{n} e^{{x_i}'\beta} - \sum_{i=1}^{n} log(y_i!) \qquad (3.9.)$$

To find ML estimation the first order partial derivative of log likelihood with respect to β equate to zero

$$\frac{\partial l(\beta; y_i)}{\partial \beta_j} = \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{n} e^{x_i{}'\beta}$$

$$= \sum_{i=1}^{n} x_i (y_i - \mu_i) = \sum_{i=1}^{n} x_i \left(y_i - e^{x_i{}'\beta}\right) = 0. \qquad (3.10.)$$

Eq. (3.10.) is the nonlinear function *of* β and cannot be directly solved. Iterative methods such as Newton - Raphson and Fisher scoring methods which noted below in Section 2.2.2., uses to solve the nonlinear equation.

3.3.2. Testing Estimated Parameters

As the member of GLMs, to test the estimated parameters of Poisson regression likelihood ratio, Wald or score tests are used which are defined by in Eq. (2.16.), (2.17.) and (2.18.) respectively. Test hypothesis is

$$H_0: \beta_i = 0, \qquad i = 1,2,\dots,n$$

$$H_1: \beta_i \neq 0, \qquad i = 1,2,\dots,n$$

The null hypothesis is rejected when the value of test statistic is a greater value than $\chi^2_{\alpha,1}$.

Likelihood ratio tests for log-linear models can easily be constructed in terms of deviances. In general, the difference in deviances between two nested models has approximately in large samples a chi-squared distribution with degrees of freedom equal to the difference in the number of parameters between the models, under the assumption that the smaller model is correct. One can also construct Wald tests, based on the fact that the maximum likelihood estimator $\hat{\beta}$ has approximately in large samples a multivariate normal distribution with mean equal to the true

parameter value β and variance-covariance matrix $var(\hat{\beta}) = X'WX$, where X is the model matrix and W is the diagonal matrix of estimation weights described earlier (Rodriguez, 2007).

3.4. Statistical Inference for Poisson Regression Model

This section includes the common measures of goodness-of-fit and residual analysis of Poisson regression model.

3.4.1. Goodness of Fit

Pearson's chi-square statistic and deviance are two common measures for goodness of fit of Poisson regression model.

The Pearson statistic for Poisson regression could be obtained by setting $\sigma_i = \mu_i$ in Eq. (2.22.)

$$P = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \tag{3.11.}$$

In large samples, the distribution of Pearson's statistic is approximately chi-squared with $df = n - p$.

For Poisson regression model chosen from general form of exponential distribution family $\hat{\theta}_i = \log(\hat{\mu}_i)$, $b(\hat{\theta}_i) = exp(\hat{\theta}_i) = \hat{\mu}_i$, for saturated model $\tilde{\theta}_i = log(y_i)$, $b(\tilde{\theta}_i) = y_i$.

Using Eq. (2.21.), deviance for Poisson regression model could be written as

$$D(y; \hat{\mu}) = 2\sum_i (y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right) - y_i + \hat{\mu}_i). \tag{3.12.}$$

Since $a(\varphi) = 1$ for Poisson, its deviance is equal to its scaled deviance (see Eq. (2.21)).

When the model with log-link contains an intercept, the likelihood equation Eq. (2.12.) implied by that parameter is $\sum_i y_i = \sum_i \hat{\mu}_i$. Deviance can be written in the form

$$D(y;\hat{\mu}) = 2\sum_i y_i \, log\left(\frac{y_i}{\hat{\mu}_i}\right).$$

Another measure for model comparison is pseudo R^2. If Eq. (2.23.) is applied to Poisson regression model, the pseudo R^2 will be

$$R^2_{pseudo} = \frac{\sum_i (y_i \log\left(\frac{\hat{\mu}_i}{\bar{y}}\right) - (y_i + \hat{\mu}_i)}{\sum_i (y_i \log\left(\frac{y_i}{\bar{y}}\right)}.$$

3.4.2. Residual Analysis

For count data, there is no one residual that has zero mean, constant variance, and symmetric distribution.

The raw residuals, Pearson residuals, and deviance residuals are commonly used for residual analysis in Poisson regression (Cameron, 2013). The general form of these residuals is shown in Section 2.5.

- The raw residual is different between the actual response and the estimated value from the model. The formula for the raw residual is

$$r_i = y_i - \hat{\mu}_i. \quad (3.13.)$$

- The Pearson residual corrects the unequal variance in the raw residuals by dividing by the standard deviation. The formula for the Pearson residual is

$$p_i = \frac{y_i - \hat{\mu}_i}{\sqrt{v(\widehat{\mu_i})}}. \quad (3.14.)$$

- The deviance residual is another popular residual. It is popular because the sum of squares of these residuals is the deviance statistic. The formula for the deviance residual is

$$r_D = \sqrt{2(y_i\, log(y_i/\hat{\mu}_i) - y_i + \hat{\mu}_i}. \qquad (3.15.)$$

4. CONTROL CHARTS

Statistical Process Control (SPC) aims at quality improvement through reduction of variation. The best-known tool of SPC is the control chart. The control chart is a successful practical technique for monitoring process measurements. A control chart is a graphical and analytic tool for monitoring process variation. There are two types of process variation (Hoblos, 2006):

- Common cause variation. The measurements from a stable process can exhibit random ups and downs. When measurements stay within acceptable limits, it is common cause variation.
- Special-cause variation. A change in the process means points below or above the control limits or measurements that trend up or down are examples of special-cause variation. Broken tool or machine, equipment degradation can be the reason of special-cause variation.

This chapter includes information about control chart, the types of control chart, and an extension of control charts to GLM. GLM based control chart with Poisson response is used as an example to this extension in this thesis.

4.1. Control Chart

A typical control chart always has three lines. All these lines are calculated by using data. A center line of control chart represents the average value of the statistical parameter feature corresponding to the in-control state. The other two lines are the *upper* and *lower* control limits. If all points fall between these two lines, then we can assume that the process is in control, otherwise, it is an evidence that the process is out-of-control (Hoblos, 2006).

A control chart is used in the following situations:

- When controlling ongoing processes by finding and correcting problems as they occur.
- When predicting the expected range of outcomes from a process.
- When determining whether a process is stable (in statistical control).
- When analyzing patterns of process variation from special causes (non-routine events) or common causes (built into the process).
- When determining whether your quality improvement project should aim to prevent specific problems or to make fundamental changes to the process.

The basic procedure to of control chart is:

1. Choosing the appropriate control chart for data;
2. Determining the appropriate time period for collecting and plotting data;
3. Collecting data, constructing chart and analyzing the data;
4. Determining "out-of-control signals" on the control chart. When one is defined, finding out the reason and how it could be corrected;
5. Correcting the data and repeating analyze. Continue the process until removal all out-of-control signals.

Control charts are quite popular for the following reasons (Montgomery, 2009):

1. Control charts are a proven technique for improving productivity
2. Control charts are effective in defect prevention
3. Control charts prevent unnecessary process adjustments
4. Control charts provide diagnostic information

5. Control charts provide information about process capability

There are some types of control charts based on data. Shewhart control chart which was first introduced by Walter Shewhart in 1924 is the most popular charts in practice. These charts are simple and quite efficient in detecting moderate to large shifts. Shewhart control charts are broadly classified into control charts for variables and control charts for attributes (Hoblos, 2006). CUSUM control charts are special types of control charts for variables. EWMA charts are the type of charts for individual observations.

4.1.1. Shewhart Control Chart for Variables

Variable charts are based on variable data that can be measured on a continuous scale. For example, weight, volume, temperature or length of stay. These can be measured to as many decimal places as necessary. Individual, average and range charts are used for variable data.

The following charts are common types of Shewhart control charts for variables:

- $\bar{x}$ and R chart -the sample *means* are plotted in order to control the mean value of a variable and the variability of a variable.
- $\bar{x}$ and s chart - the sample *standard deviations* are plotted in order to control the variability of a variable.
- s^2 chart - the sample *variances* are plotted in order to control the variability of a variable.

Directly obtained process mean and variance makes all control charts for variables more informative. The variables charts will often signal a process problem

before the process has produced nonconforming items and it provides visual means to directly assess process performance regardless of the specifications.

Assume that a quality characteristic is normally distributed with mean μ and standard derivation σ, where both μ and σ are known. If $x_1, x_2, \dots, x_n$ is a sample of size n, the average of this sample is

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Since $\bar{x}$ follow the normal distribution with mean μ and standard derivation $\sigma_{\bar{x}} = \sigma/\sqrt{n}$, the probability is $1-\alpha$ of any sample mean will change between

$$\mu \pm Z_{\alpha/2}\sigma_{\bar{x}} = \mu \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}. \qquad (4.1.)$$

If μ and σ are known, Eq. (4.1.) could be used as an upper and lower control limit of the control charts for sample means.

If μ and σ are unknown, they must be estimated from subgroups taken when the process is thought to be in control.

Assume that there are m samples and each of them contains n observations on the quality characteristic. Let $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m$ be an average of each sample. The process average is the grand average

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_m}{m}. \qquad (4.2.)$$

If $x_1, x_2, \dots, x_n$ is a sample of size n, the range of the sample is

$$R = x_{max} - x_{min}.$$

Let $R_1, R_2, \dots, R_m$ be the ranges of the samples and the average range is

$$\bar{R} = \frac{R_1 + R_2 + \cdots + R_m}{m}. \qquad (4.3.)$$

There is $W = R/\sigma$, a random variable which is called relative range. The mean and the standard deviation of W is d_2 and d_3 respectively (d_2 and d_3 are tabulated values shown in APPENDIX 1). The estimation of σ is

$$\hat{\sigma} = \bar{R} \ / d_2 \tag{4.4.}$$

If an estimator of μ is used and $\bar{R}/d_2$ as estimator of σ than the control limits for $\bar{x}$ chart are

$$\begin{aligned} UCL &= \bar{\bar{x}} + \frac{3}{d_2/\sqrt{n}}\bar{R} \\ Center\ Line &= \bar{\bar{x}} \\ LCL &= \bar{\bar{x}} - \frac{3}{d_2/\sqrt{n}}\bar{R}\,. \end{aligned} \tag{4.5.}$$

Estimation of the standard derivation of *R* could be founded by using relative range $W = R/\sigma$. Since $R = W\sigma$ the standard derivation of *R* is

$$\sigma_R = d_3\sigma.$$

Since σ is unknown, the σ_R could be estimated by following way

$$\hat{\sigma}_R = d_3 \frac{\bar{R}}{d_2}. \tag{4.6.}$$

The control limits for *R* chart are

$$\begin{aligned} UCL &= \bar{R} + 3\hat{\sigma}_R = \bar{R} + 3d_3 \frac{\bar{R}}{d_2} \\ Center\ line &= \bar{R} \\ LCL &= \bar{R} - 3\hat{\sigma}_R = \bar{R} - 3d_3 \frac{\bar{R}}{d_2}. \end{aligned} \tag{4.7.}$$

Defining $A_2 = 3/(d_2/\sqrt{n})$, $D_3 = 1 - 3(d_3/d_2)$ and $D_4 = 1 - 3(d_3/d_2)$, the control limits for $\bar{x}$ and *R* chart in Eq. (4.5.) and Eq. (4.7.) could be written in the following form:

Control limits for $\bar{x}$ chart

$$\begin{aligned} UCL &= \bar{\bar{x}} + A_2\bar{R} \\ Center\ Line &= \bar{\bar{x}} \\ LCL &= \bar{\bar{x}} - A_2\bar{R}\,. \end{aligned} \tag{4.8.}$$

and control limits for *R* Chart are

$$UCL = D_4\bar{R}$$
$$Center\ Line = \bar{R} \qquad (4.9.)$$
$$LCL = D_3\bar{R}$$

The values of A_2, D_3 and D_4 are constants (APPENDIX 1).

$\bar{x}$ and *s* control charts, where *s* is the sample standard derivation are preferable in following cases

1. The sample size is large, $n > 10$ *or* 12.
2. The sample size *n* is variable

The sample variance is

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}. \qquad (4.10.)$$

The control limits for *s* control chart are

$$UCL = B_6\sigma$$
$$Center\ Line = c_4\sigma \qquad (4.11.)$$
$$LCL = B_5\sigma$$

where B_5 and B_6 are constants (APPENDIX 1) and equals $c_4 - 3\sqrt{1-c_4^2}$ and $c_4 + 3\sqrt{1-c_4^2}$, respectively. c_4 is a constant (APPENDIX 1) that depends on sample size *n*. If no standard is given for σ, then it must be estimated by analyzing past data.

Assume that *m* samples are available, each of size *n* and let s_i be the standard derivation of the *i*th sample. The average of the m standard derivation is

$$\bar{s} = \frac{1}{m}\sum_{i=1}^{m} s_i.$$

The statistic $\bar{s}/c_4$ is unbiased estimator of σ. Therefore, the parameters of s chart would be

$$UCL = \bar{s} + 3\,\frac{\bar{s}}{c_4}\sqrt{1-c_4^2}$$

$$Center\ Line = \bar{s} \qquad (4.12.)$$

$$UCL = \bar{s} + 3\,\frac{\bar{s}}{c_4}\sqrt{1-c_4^2}\,.$$

Let B_3 and B_4 (see APPENDIX 1) are constants and equals to $1-\frac{3}{c_4}\sqrt{1-c_4^2}$ and $1+\frac{3}{c_4}\sqrt{1-c_4^2}$ respectively, then the control limits for *s* chart in Eq. (4.12.) would be

$$UCL = B_4\bar{s}$$

$$Center\ Line = \bar{s} \qquad (4.13.)$$

$$LCL = B_3\bar{s}\,.$$

The s^2 control chart is directly based on sample variance. The control limits for s^2 control chart are

$$UCL = \frac{\bar{s}^2}{n-1}\,\chi^2_{1-(\alpha/2),n-1}$$

$$Center\ Line = \bar{s}^2 \qquad (4.14.)$$

$$LCL = \frac{\bar{s}^2}{n-1}\,\chi^2_{1-(\alpha/2),n-1}$$

where $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1-(\alpha/2),n-1}$ are upper and lower $\alpha/2$ percentage points of the chi-squared distribution with *df=n-1*, $\bar{s}^2$ is an average sample variance.

4.1.2. Shewhart Control Chart for Attributes

Attribute charts are based on data that can be grouped and counted as present or not. These charts are also called count charts, and attribute data are also known as discrete data. Attributes control charts are graphs that display the value of a process variable over time which includes points, a centerline which represents the average of the monitored variable and upper and lower control limits. There are several types of attributes control charts Woodall (1997).

4.1.2.1. C Chart

A *C* chart is an attribute control chart that displays how the number of defects or nonconformities, for a process is changing over time. The number of defects is collected for the group of units or just one individual unit on which defect counts are performed in each subgroup. A *C* chart is based on Poisson distribution.

The control limits for *C* chart are

$$\begin{aligned} UCL &= c + 3\sqrt{c} \\ Center\ line &= c \\ LCL &= c - 3\sqrt{c} \end{aligned} \tag{4.15.}$$

where $c > 0$ is a parameter of Poisson distribution. For the negative value of LCL, LCL sets equal to the zero.

4.1.2.2. U Chart

A *U* chart (u is for the unit) is an attribute control chart that displays how the frequency of defects, or nonconformities, is changing over time for a process. A *U* chart is an indicator of the consistency and predictability of the level of defects in the process. A *U* chart is appropriate when the units for detecting defects vary from subgroup to subgroup. Since this can be seen in the shifting *UCL* and *LCL* lines that depend on the size of the subgroup. This chart shows the number of defects per unit. A *U* chart is based on Poisson distribution.

The control limits for *U* charts are

$$UCL = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}}$$

$$Center\ line = \bar{u} \qquad (4.16.)$$

$$LCL = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}}$$

where $\bar{u}$ is the observed average number of defects per unit in a set of data.

4.1.2.3. Np Chart

Np charts also are known as a number of rejects chart for constant subgroup size. The *Np* chart is an attribute control chart that displays changes in the number of defective products, rejects or unacceptable outcomes. It is an indicator of the consistency and predictability of the level of defects in the process. The *Np* chart is only valid as long as data collected from subgroups that are the same size. A *Np* chart is based on binomial distribution.

The control limits for the *Np* chart are

$$UCL = n\bar{p} + 3\sqrt{n\bar{p}(1 - n\bar{p})}$$

$$Center\ Line = n\bar{p} \qquad (4.17.)$$

$$LCL = n\bar{p} - 3\sqrt{n\bar{p}(1 - n\bar{p})}$$

where $\bar{p}$ is the estimate of the long-term process mean established during control-chart setup.

4.1.2.4. P Chart

A *P* chart is an attribute control chart that displays changes in the proportion of defective products, rejects or unacceptable outcomes. It is an indicator of the consistency and predictability of the level of defects in the process. Since a *P* chart is used when the subgroup size varies, the chart plots the proportion or fraction of items

rejected. This is indicated by the shifting *UCL* and *LCL* lines that depend on the size of the samples on *P* chart. The *P* chart is based on binominal distribution.

The control limits for *P* chart are

$$UCL = \bar{p} + 3\sqrt{\bar{p}(1-n\bar{p})/n}$$
$$CL = \bar{p} \qquad (4.18.)$$
$$LCL = \bar{p} - 3\sqrt{\bar{p}(1-n\bar{p})/n}$$

where $\bar{p}$ is average of defects and *n* is the number of items in a given subgroup.

4.1.3. CUSUM Control Chart

The CUSUM chart was first proposed by Page (1954) to monitor a normal mean. Most of its popularization may be attributable to Lucas (1976). Unlike the Shewhart chart, CUSUM chart utilizes the past information available from previously plotted points, so has a greater sensitivity for detecting small shifts. Brook et al. (1972) and Lucas (1985) purposed CUSUM charts for attributes with a Poisson model for the counts of defects.

As noted below the CUSUM control chart is more efficient in detecting small shifts in the mean of a process. It directly incorporates all the information in the sequence of sample values by plotting the cumulative sums of the derivations of the sample values from a target value. Shewhart charts detect large and sudden changes in measurement, as a two or three sigma shift and they are not effective at detecting smaller changes, such as a one sigma shift like CUSUM control chart.

Suppose that there are $n \geq 1$ samples and $\bar{x}_j$ is the average of the *j*th sample. If μ_0 is the target for the process mean, the CUSUM control chart is formed by plotting the quantity

$$C_i = \sum_{j=1}^{i} (\bar{x}_j - \mu_0) \qquad (4.19.)$$

If process is in control at the target value μ_0 Eq. (4.19.) is a random walk with zero mean. If $\mu_0 < \mu_1$ than positive drift will develop in cumulative sum C_i, if $\mu_0 > \mu_1$ than negative drift will develop in cumulative sum C_i.

There are two forms of representing CUSUM

- Tabular (or algorithmic) CUSUM
- V-mask CUSUM

4.1.3.1. Tabular CUSUM

Let x_i be the ith observation on the process. When process is in control x_i has a normal distribution with mean μ_0 and standard derivation σ. Assume that the σ is known or could be estimated.

The tabular CUSUM works by accumulating derivations from μ_0 that are above target statistic C^+ and below target statistic C^-. These statistics are called one-sided upper and lower CUSUMs respectively and computed as follows:

$$C^+ = max[0, x_i - (\mu_0 + K) + C_{i-1}^+]$$
$$C^- = max[0, (\mu_0 - K) - x_i + C_{i-1}^-] \quad (4.20.)$$

where the starting values are $C_0^+ = C_0^- = 0$. K is the "reference" value specified for plotted CUSUM line. The CUSUM values are accumulated only when the deviation from the target value exceeds specified K value. K is often chosen about halfway between target μ_0 and out-of-control value μ_1. Than K is

$$K = \frac{|\mu_1 - \mu_0|}{2}\ \sigma. \quad (4.21.)$$

The tabular CUSUM also indicates when the shift occurred. The counter N^+ records the number of consecutive periods in upper-side CUSUM C_i^+ rose above the value of zero and N^- records the number of consecutive periods in downside C_i^- fall below the value of zero.

The graphical display of tabular CUSUM is called CUSUM status charts. They are constructed by plotting C_i^+ and C_i^- versus the sample number.

H is the "Decision Interval" which acts as a control limit

$$H = h\sigma$$

where σ is standard derivation of the sample variable used in forming the CUSUM and *h* is the decision parameter. *h* must be chosen to give the desired in-control ARL_0 performance. Using $h = 4$ or $h = 5$ will generally provide a CUSUM that has good ARL properties against a shift about 1σ in the process mean (Montgomery, 2009).

4.1.3.2. V-Mask CUSUM

V-Mask is another way for representing CUSUM. This procedure is proposed by Bernard (1959).

V-mask is applied to successive values of the CUSUM statistic

$$C_i = \sum_{j=1}^{i} y_j = y_i + C_{i-1} \qquad (4.22.)$$

where y_i is the standardized observation $y_i = (x_i - \mu_0)/\sigma$.

The decision procedure consists of placing the V-mask on the cumulative sum control chart with the point *O* on the last value of C_i and the line *OP* parallel to the horizontal axis. If all the previous cumulative sums, C_1, C_2, ... , C_i lie within the two arms of the V-mask, the process is in control. However, if any of the cumulative sums lie outside the arms of the mask, the process is considered to be out of control. In actual use, the V-mask would be applied to each new point on the CUSUM chart as soon as it was plotted, and the arms are assumed to extend backward to the origin. The performance of the V-mask is determined by the lead distance *d* and the angle θ shown in Figure 4.1.

The results of CUSUM control charts determined by parameters of V-mask. The parameters of V-mask is the distance *d* which is defined as

$$d = \left(\frac{2}{\delta^2}\right)\ln\left(\frac{1-\beta}{\alpha}\right)$$

and the angle θ

$$\theta = tan^{-1}\left(\frac{\Delta}{2A}\right)$$

where 2α is the greatest allowable probability of a signal when the process mean equal to the target value, δ is the amount of shift (normalized to σ) that we wish to detect, β is the probability of not detecting a shift of size δ and A is the horizontal distance on the V-mask plot between successive points in terms of unit distance on the vertical scale.

When β is small, then

$$d \simeq -2\frac{\ln(\alpha)}{\delta}.$$

The tabular CUSUM and V-mask CUSUM are equivalent in the case

$$k = Atan\theta$$

and

$$h = Adtan\theta = dk.$$

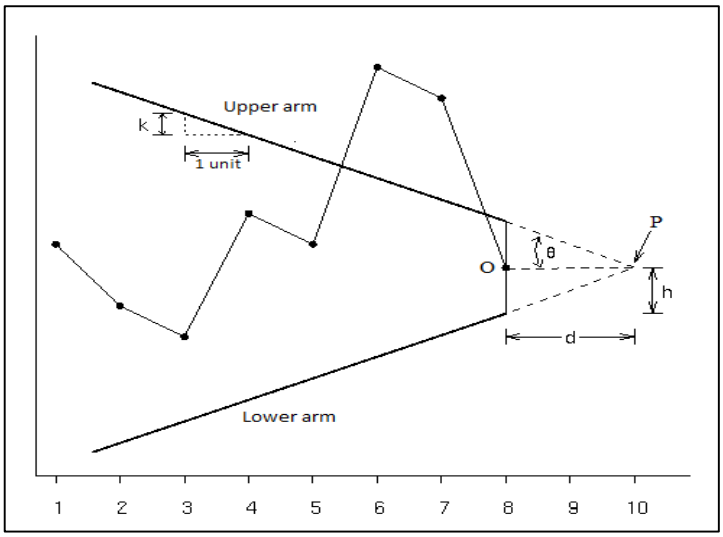


Figure 4.1. Typical V mask CUSUM (Montgomery, 2009)

The CUSUM control chart is good at detecting out-of-control observations in small shift sizes, but V-mask CUSUM control chart have some disadvantages:

1. The head start feature, which is very useful in practice, cannot be implemented with the V-mask.

2. It is sometimes difficult to determine how far backward the arms of the V-mask should extend, thereby making interpretation difficult for the practitioner.

4.1.4. EWMA Control Chart

EWMA control charts were first proposed by Roberts (1959). EWMA and CUSUM control charts are competing charts to detect small shifts in the mean of normally distributed observations in research by Roberts (1966). EWMA control charts also accumulate current and past information from the observations and more effective in detecting small shifts in the process mean. Box et al. (1974), Robinson et al. (1978), Hunter (1986), Crowder (1987), and Waldmann (1986) have evaluated the properties of the EWMA control charts in their works analytically. Crowder (1987) and Lucas and Saccucci (1990) provide a significant contribution to the design of an optimal EWMA control chart.

The EWMA control charts are also known as the geometric moving average control charts. These charts are used to monitor the mean of a process based on samples taken from the process at given times and typically used for individual observations. EWMA control charts are simpler to set and use.

For EWMA control chart, the decision depends on the EWMA statistic. EWMA is a statistic for monitoring the process that averages the data in a way that gives less and less weight to data as they are further removed in time.

The EWMA statistic is calculated as

$$z_i = \lambda x_i + (1 - \lambda) z_{i-1} \qquad (4.23.)$$

where $0 < \lambda \leq 1$ is a constant.

The starting value when *i=1* is $z_0 = \mu_0$.

EWMA control chart is a weighted average of all past and current observations. It makes this control chart sensitive to normality assumptions and ideal for individual observations.

If observations x_i are independent random variables with variance σ^2, the variance of z_i is

$$var\,(z_i) = \sigma^2 \left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right). \tag{4.24.}$$

The EWMA control chart is constructed by plotting z_i against the sample number *i* or time. The control limits for EWMA control chart are

$$\begin{aligned} UCL &= \mu_0 + L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)(1-(1-\lambda)^{2i})} \\ Center\ Line &= \mu_0 \\ LCL &= \mu_0 - L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)(1-(1-\lambda)^{2i})} \end{aligned} \tag{4.25.}$$

where *L* is the width of the control limits.

When *i* in Eq. (4.24.) gets larger $\left(1-(1-\lambda)^{2i}\right)$ approaches unity. In other words, after the EWMA control chart has been running for several time periods, control limits will approach steady-state values given by

$$\begin{aligned} UCL &= \mu_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}} \\ Center\ Line &= \mu_0 \\ LCL &= \mu_0 - L\sigma\sqrt{\frac{\lambda}{2-\lambda}}. \end{aligned} \tag{4.26.}$$

The EWMA control charts have some attractive properties (Cisar et al., 2011):

1. Unlike $\bar{x}$ and R control charts, all of the data collected over time may be used to determine the control status of a process.
2. Like the CUSUM, the EWMA utilizes all previous observations, but the weight attached to data exponentially decreases as the observations become older and older.
3. The EWMA is often superior to the CUSUM charting technique due to the fact that it detects larger shifts better.
4. The EWMA schemes may be applied for monitoring standard deviations in addition to the process mean.
5. The EWMA schemes can be used to forecast values of a process mean.
6. The EWMA methodology is not sensitive to normality assumptions.

4.2. Generalized Linear Model Based Control Charts

4.2.1. GLM Based Control Chart

A regression based control chart which is the combination of the control chart techniques and simple linear regression models was first proposed by Mandel (1969). The basic idea is to control varying rather than a constant mean. Mandel (1969) used regression control chart to monitor the variety of postal management problems. The modified regression control chart is also used when the process exhibits tool wear (Montgomery, 2009). Mandel (1969) also devised a simplification of the regression control chart. Zhang (1984) adopted Mandel's idea of a residual control chart for statistical process control data to the cause selecting chart which is a type of a regression based control chart, is constructed for an outgoing quality characteristic only after it has been adjusted for the effect of incoming quality characteristic. A regression adjustment control chart, developed by Hawkins (1991; 1993), monitors the residuals from the process variable of interest when that variable is regressed on all the others. The residuals should be independently distributed over *N(0,1)* if

process parameters are well estimated. When applying traditional charts to residuals, they are assumed to be statistically uncorrelated. The main advantage of a residual chart is that it can be applied to any autocorrelated data whether the process is stationary or not (Sauza et al., 2015).

Loredo (2002) applied the individual measurement chart for residuals of multiple linear regression models and Shu (2004) proposed the EWMAREG chart, which basically consists in monitoring the standardized residuals of the regression model with EWMA control chart. Control charts proposed by Haworth (1996), Loredo (2002) and Shu (2004) present the further advantage of preserving the temporal order of data, which makes the application of these procedures and the interpretation of results easier if compared to the method presented by Mandel (1969).

The difference of regression based control charts from traditional control charts is (Mandel, 1969):

- It is designed to control a varying (rather than a constant) average.
- The control limit lines are parallel to the regression line rather than the horizontal line.
- The computations here are much more complex.
- It is appropriate for use in more complex situations

The success of the regression based control charts method warrants (Hawkins, 1993):

- A review of statistical and technical analyses required to establish and use this chart
- Giving illustrative applications to a variety of management problems.

GLM based control charts have been proposed for monitoring multiple discrete counts (Skinner et al., 2003).

The quality characteristics are often the combination of normal and non-normal distributed variables. When the variables follow Gamma, Normal or Poisson distributions, GLM based control method must be used. The GLM based control charts are suitable to detect changes in the mean of the quality characteristics.

The GLM can be used to enhance the ability of regression based control chart to perform effectively in an environment where input and output variables are both normally or non-normally distributed.

The steps to implement GLM based control charts are:

Step1. Obtain the data to process

Step2. Fit the GLM to the response and predicted variables from in-control data

Step3. Obtain the residuals for in-control process data

Step4. Obtain the control limits to control charts

Step5. Monitor the residuals

4.2.2. GLM Based Control Charts with Poisson Response

When the response variables follow the Poisson distribution GLM based control chart could be used instead of the Shewhart attribute charts for count data. Poisson Shewhart, Poisson CUSUM, Poisson EWMA control charts are basic methods to monitor the Poisson distributed observations. For monitoring process deviance residuals are obtained from the regression model and used as a control chart statistics. The deviance residuals are commonly used residual type in this process since deviance residuals are asymptotically normally distributed.

4.2.2.1. Poisson Shewhart Control Chart

Applying the residuals to control charts for process control the control limits for Shewhart control chart can be constructed by using the deviance residuals for Poisson distributed data. The deviance residual for Poisson regression model is defined in Eq. (3.12.).

The control limits based on deviance residuals are

$$\begin{aligned} UCL &= \mu_r + k\sqrt{var(\bar{r}_D)} \\ Center\ Line &= \mu_r \\ LCL &= \mu_r - k\sqrt{var(\bar{r}_D)} \end{aligned} \tag{4.27.}$$

where μ_r is the mean of deviance residual.

The constant k defines the width for control limits based on ARL until the false alarm is verified. For measurement, the control limits based on the asymptotic performance of control chart ARL.

4.2.2.2. Poisson CUSUM Control Chart

To monitor the Poisson distributed data CUSUM control charts are more effective to detect small shifts than the Shewhart control charts.

If we apply residuals to CUSUM control chart statistic changes to:

$$C_i = r_{D\,i} - \mu_r + C_{i-1} \tag{4.28.}$$

where r_D is deviance residual obtained from count data and μ_r is the mean of the deviance residuals. For one sided CUSUM we use Eq. (4.20.) by replacing observations with deviance residuals.

Decision intervals which act as control limits are

$$\begin{aligned} H^+ &= h\sigma \\ H^- &= -h\sigma \end{aligned} \tag{4.29.}$$

where σ is standard derivation of the deviance residuals and h is the decision parameter.

The reference value in normal mean case is

$$k = k^- = k^+ = \frac{\mu_1 - \mu_0}{2}\sigma.$$

It is common practice and sensible to round the value of k to some more convenient value. There are two common reasons for this. Both of them are matters of convenience. By limiting the number of attainable values of the CUSUM, the speed of the calculation of ARL could be speeded up. In this case, k^+ and k^- values changes to

$$k^+ = \frac{\mu_1 - \mu_0}{ln\,\mu_1 - ln\,\mu_0} \; and \; k^- = \frac{\mu_0 - \mu_1}{\ln \mu_0 - ln\,\mu_1}.$$

If C^+ exceeds the decision interval H^+, then the chart signals than an increase in the mean count have occurred and if C^- exceeds H^-, then the chart signals than a decrease in the mean count have occurred.

4.2.2.3. Poisson EWMA Control Chart

Residuals could be also applied to EWMA control charts. Defining $z_0 = \mu_r$, and EWMA statistic is

$$z_i = \lambda r_D + (1 - \lambda) z_{i-1} \qquad (4.30.)$$

where $0 < \lambda \leq 1$ is a constant and r_D is deviance residual obtained from count data.

The control limits and the center line for the EWMA control chart is computed by

$$\begin{aligned} UCL &= \mu_r + L\sigma\sqrt{(\lambda/(2-\lambda))var(r_D)} \\ Center\ Line &= \mu_r \\ LCL &= \mu_r - L\sigma\sqrt{(\lambda/(2-\lambda))var(r_D)} \end{aligned} \qquad (4.31.)$$

where L is the width of the control limits and μ_r is the mean of the deviance residuals.

From the definition of EWMA in Eq. (4.30.), the mean is $E(z_i) = \mu_0$. When $\mu = \mu_0$ the process is under control. When the mean shifts to another value μ_1, $\mu_1 > \mu_0$ or $\mu_1 < \mu_0$ the process is out of the control. In both cases the quick detection of shifts is desirable.

The variance is $var\,(z_i) = \left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)$. For large values of *i*, variance is approximately

$$var\,(z_i) \approx \frac{\lambda\mu_0}{2-\lambda}\,. \tag{4.32.}$$

The control limits based on the asymptotic variance in Eq. (4.32.) are

$$\begin{aligned} UCL &= \mu_r + A_U^r\sqrt{\left(\frac{\lambda}{2-\lambda}\right)(1-(1-\lambda)^{2i})} \\ Center\ Line &= \mu_0 \\ LCL &= \mu_r - A_L^r\sqrt{\left(\frac{\lambda}{2-\lambda}\right)(1-(1-\lambda)^{2i})} \end{aligned} \tag{4.33.}$$

where A_U and A_L are upper and lower control limit factors and often chosen to be equal ($A_U^r = A_L^r = A^r$).

4.3. Measuring Performance of Control Chart

There are several methods for measuring the performance of control charts. Average of run length and standard deviation of run length, the average time for the signal are three of them.

The techniques to detect the shift in the mean or in the variance of the process are under an assumption that the observations are independent and normal distributed. But the data are often non-normal and correlated. The performance of a control chart is usually evaluated by analyzing the properties of its run length (RL) distribution. RL can be defined as the number of points which are the control chart statistics plotted on the chart until the first out-of-control signal is generated. The

mean of the RL distribution is the ARL, which is actually the average number of observations needed for a control chart to the signal. Along with the ARL, the standard deviation of the run length (SDRL) is computed. Alternatively, the ARL is expressed as the average number of observations to signal (ANOS).

Another way to measure of performance of control charts is the average time to signal (ATS). ATS is a measure similar to the ARL and the average time needed for a control chart to signal and it is actually a product of the ARL and the sampling interval used in the case of fixed sampling.

ARL is used for performance comparison of control charts in this thesis.

4.3.1. Average Run Length

ARL is the most common used metric for evaluating the performance of a control chart. ARL can be defined as the expected number of points plotted on a control chart until a point indicates an out-of-control situation. In other words, it is the mean of RL distribution.

The ARL of the chart when the process is in-control is denoted by ARL_0. Since an in-control state of the process is desired to be long, it is advantageous for a process to have large ARL_0 values.

When the process shifts or in other words the process is out-of-control, the early detection of the shift is essential. The number of points plotted needed to detect a shift is denoted by ARL_1. A chart with a small ARL_1 is preferred. Moreover, a control chart with a smaller ARL_1 value for a particular shift magnitude is generally preferred to the ones with larger ARL_1 values for the same shift (Testik et al. 2006).

ARL is very important in examining the performance of the control chart because it indicates how early each of our methods is able to catch the misbehaving source. The earliest the technique detects the misbehaving source, the better the method is.

The ARL is defined as

$$ARL = \frac{1}{probability} \tag{4.34.}$$

where $probability$ describes the probability that any point exceeds the control limits (Montgomery, 2009). This equation can be used to evaluate the performance of the control chart.

4.3.1.1. ARL for Shewhart Control Chart

The ARL value for any Shewhart control chart is

$$ARL_0 = \frac{1}{\alpha} \tag{4.35.}$$

where $\alpha = 1 - P\{LCL < \bar{x} < UCL|\mu = \mu_0\}$, when process in-control and

$$ARL_1 = \frac{1}{1-\beta} \tag{4.36.}$$

when the process is out-of-control where β is the risk and probability of not detecting this shift on the first subsequent sample. For the $\bar{x}$ control chart with known and constant standard derivation σ, if the mean shifts from the in-control value μ_0 to value $\mu_1 = \mu_0 + k\sigma$ the β- risk is

$$\beta = P\{LCL < \bar{x} < UCL|\mu = \mu_1 = \mu_0 + k\sigma\}. \tag{4.37.}$$

The probability of the detected shift on the *j*th sample subsequent is $\beta^{j-1}(1-\beta)$. And the expected number of samples taken before the shift is detected as

$$ARL_1 = \sum_{j=1}^{\infty} j\beta^{j-1}(1-\beta) = \frac{1}{1-\beta}.$$

4.3.1.2. ARL for CUSUM and EWMA Control Chart

There are several techniques for calculating ARL of a CUSUM.

ARL for one-sided CUSUM with parameters *k* and *h* is

$$ARL = \begin{cases} (\exp(-2\Delta b) + 2\Delta b - 1)/\ 2\Delta^2, & \Delta = 0 \\ b^2, & \Delta \neq 0 \end{cases} \quad (4.38.)$$

where $\Delta = \delta^* - k$ for the upper one-sided CUSUM C_i^+ and $\Delta = -\delta^* - k$ for the lower one-sided CUSUM C_i^-. $b = h + 1.166$ and $\delta^* = \frac{\mu_1 - \mu_0}{\sigma}$.

The δ^* is the shift in the mean, in the units of σ, for which the ARL is to be calculated (SAS, 2012).

If $\delta^* = 0$ than ARL_0 is calculated by using Eq. (4.38.), if $\delta^* \neq 0$ the value of ARL_1 would be calculated corresponding to a shift of size δ^*.

Using two one-sided statistics ARL^+ and ARL^-, ARL of two-sided CUSUM could be obtained as

$$\frac{1}{ARL} = \frac{1}{ARL^+} + \frac{1}{ARL^-} \quad (4.39.)$$

The ARL performance for the EWMA control chart is closely approximated CUSUM ARL performance and depends on chosen value of *L* and λ (Montgomery, 2009). Specifying the desired in-control and out-of-control ARLs and the magnitude of the process shift that is anticipated, and then selecting the combination of *L* and λ that provide the desired ARL performance is the optimal design procedure.

5. APPLICATION ILLUSTRATION

Empirical and real case study have been illustrated in this chapter. The results of studies are shown in tables and figures.

5.1. Empirical Study

The performance of control charts has been developed by the simulations to calculate run length of control charts using MATLAB R2015a. Deviance residuals which are obtained from Poisson regression model are used as statistics for plotting control chart. Pierce and Shafer (1986) recommend the deviance residuals beside other residuals in GLM. Deviance residuals are asymptotically normally distributed and authors noted that the deviance residuals are normally distributed for remarkably small sample sizes.

The 1000 replications for each trial with n=100 observations to analyze the performance of control charts with Poisson response are simulated.

The steps to apply model based control charts are summarized here:

1. Generate *Nprelim* (in control preliminary samples) samples (x)
2. Choose arbitrary constants for regression coefficients
3. Obtain the mean of the response variable by using the inverse of link function for exponential family
4. Generate response variables with obtained mean
5. Obtain deviance residuals for in-control process data
6. Calculate in-control average and standard deviation based on the *NPrelim* simulated in-control deviance residuals

7. Determine the upper and lower control limits of each control chart (Shewhart, CUSUM, and EWMA) to achieve the desired ARL_0
8. Calculate out-of-control residuals under several types of shifts
9. Repeat step 1 through step 8 until the number of simulations ($Nsim$) is reached. The result of $Nsim$ is run length
10. Find an average of run lengths which is equal to ARL.

The variables x_1 and x_2 are assumed independent and identically normal distributed with zero mean and unit variance. The sample size for simulation is chosen as 100. The μ_i represents the mean given an observation vector $[x_{i1}, x_{i2}]$, where for each $i = 1, \dots, n$, the value y_i is simulated using the Poisson distribution parameter μ_i given by Eq. (3.6). The deviance residuals are calculated from Eq. (3.15) and are used as control chart statistics. In-control mean and in-control standard derivation are calculated as a mean of means and mean of standard derivations of residuals. By using the in-control mean and in-control standard derivation, the control limits for Shewhart, CUSUM and EWMA control charts are calculated from Eq. (4.27.), (4.29.) and (4.31.). The k value for Shewhart control chart, the h value for CUSUM control chart and the λ and A^r values for EWMA control chart are determined so as to achieve $ARL_0 = 200$.

The new set of observations is generated under shifts and the observed values are used to predict the mean values. The out-of-control deviance residuals are computed from the predicted values. The new observations are generated until all control charts are out-of-control.

The control limits of Shewhart control chart are determined to achieve $ARL_0=200$. As a result, the upper and the lower control limits for Shewhart control

chart are 2.6014 and -2.8729, for CUSUM control chart are 2.9054 and -2.9054 and for EWMA control chart are 2.4220 and -2.6949, respectively. EWMA control chart has a wide distance between control limits, whereas CUSUM control chart has a narrow distance between control limits. The simulation repeated 1000 times under several types of shifts.

Two types of shifts are considered in simulations:

- A multiplicative shift, which is a change in any regression coefficient or coefficients for shifted response $\mu_1 = g^{-1}(x'(\beta + v))$
- An additive shift which is a shift in mean for shifted response $\mu_1 = g^{-1}(x'\beta) + v$

The regression parameters are chosen as $\beta_0 = 1$, $\beta_1 = 2$ and $\beta_2 = 1$, and the shift is chosen as $v = \delta * \sigma$ with the shift size δ=0.5, 1, 1.5, 2, 2.5, 3 and the standard deviation of in-control observations σ.

The process is repeating until the number of simulations is reached.

5.1.1. Performance Analysis of Control Charts

The performance analysis of the control charts is examined by their ARL value.

The number of points which are the control chart statistics plotted on the chart until the first out-of-control signal is generated is the run length. The average of the run lengths is ARL.

5.1.1.1. Changes in the Regression Coefficients

In this section, the performance of control charts under shifts in regression coefficients is illustrated.

The results of the simulation experiments where the parameter β_i is shifted to $\beta_i \pm v, i = 0,1,2$ are summarized in Table 5.1., 5.2., and 5.3., respectively.

Table 5.1. The *ARL* performance of control charts under shift in $\beta_0 \pm \delta\sigma$

Chart type	**Shift type**	**Shift size, δ**					
		0.5	**1**	**1.5**	**2**	**2.5**	**3**
Shewhart	$\beta_0 + \delta\sigma$	82.39	80.54	82.1	83.95	84.48	74.75
	$\beta_0 - \delta\sigma$	77.93	76.32	80.72	73.33	79.79	75.39
CUSUM	$\beta_0 + \delta\sigma$	5.2	25.5	17.4	5.6	6	24.3
	$\beta_0 - \delta\sigma$	50.7	10	23.3	12.8	20	31.9
EWMA	$\beta_0 + \delta\sigma$	98.64	99.12	98.72	98.34	99.6	98.95
	$\beta_0 - \delta\sigma$	97.11	96.88	98.75	99.34	97.32	97.75

Table 5.2. The *ARL* performance of control charts under shift in $\beta_1 \pm \delta\sigma$

Chart type	**Shift type**	**Shift size, δ**					
		0.5	**1**	**1.5**	**2**	**2.5**	**3**
Shewhart	$\beta_1 + \delta\sigma$	78.03	78.81	75.75	78.54	84.68	73.59
	$\beta_1 - \delta\sigma$	80.31	75.11	80.38	79.16	78.18	78.63
CUSUM	$\beta_1 + \delta\sigma$	41.3	21.2	12.6	10.3	9.6	13
	$\beta_1 - \delta\sigma$	53.3	24.5	48.6	34.5	21.6	8
EWMA	$\beta_1 + \delta\sigma$	99.47	99.07	98.38	98.12	97.9	98.57
	$\beta_1 - \delta\sigma$	99.16	97.02	99.7	98.73	99.22	99.02

Table 5.3. The *ARL* performance of control charts under shift in $\beta_2 \pm \delta\sigma$

Chart type	**Shift type**	**Shift size, δ**					
		0.5	**1**	**1.5**	**2**	**2.5**	**3**
Shewhart	$\beta_2 + \delta\sigma$	82.54	86.2	78.58	82.31	79.9	74.78
	$\beta_2 - \delta\sigma$	75.55	78.9	79.39	84.07	74.99	84.65
CUSUM	$\beta_2 + \delta\sigma$	32.1	26.5	17.6	9.8	6.8	8
	$\beta_2 - \delta\sigma$	47	41.4	25.2	14	8.5	1
EWMA	$\beta_2 + \delta\sigma$	95.71	97.47	99.5	99.17	99.12	100
	$\beta_2 - \delta\sigma$	100	98.47	99.73	98.75	98.01	99.18

As shown in Table 5.1., 5.2., and 5.3. The conclusions obtained from Tables are as follows:

- CUSUM control chart outperforms under all shifts considered and shows small ARL values. In other words, the CUSUM control chart detects the out-of-control observation quickly than other control charts especially, when β_2 is shifted positively and β_1 is shifted negatively.
- The EWMA control chart shows the worst result among others, especially when β_2 is shifted positively and negatively.
- Shewhart control chart partially shows a good performance when β_1 or β_2 is shifted.
- The CUSUM control chart is sensitive to changes, the ARL values are sharply decreased when shift size increased. The Shewhart control chart is also showed the small changing depend on the shift size. The EWMA control chart not shows serious changes depend on the shift size.

The summary of simulation results where the two of the parameters are shifted as $\delta\sigma$ at the same time are shown in Table 5.4., 5.5., and 5.6., respectively.

Table 5.4. The ARL performance of control charts under shift in $\beta_0 \pm \delta\sigma$ and $\beta_1 \pm \delta\sigma$

Chart type	**Shift type**	**Shift size, δ**					
		0.5	**1**	**1.5**	**2**	**2.5**	**3**
Shewhart	$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma$	99.53	89.03	88.89	88.8	87.25	84.45
	$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma$	98.38	77.49	76.7	75.31	71.59	69.62
	$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma$	99.63	85.89	83.79	81.05	80.35	80.22
	$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma$	99.12	81.88	77.22	75.83	75.03	73.62
CUSUM	$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma$	35.3	34.5	18.1	13	12.3	10.7
	$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma$	71.8	61.1	50.8	47.8	41.2	32
	$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma$	67	39.3	35.9	34.7	31.3	23.8
	$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma$	47.7	39.3	3.2	2.8	1.3	0.19
EWMA	$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma$	96.02	79.04	30.52	15.97	9.71	1.44
	$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma$	99.33	98.19	99.5	97.2	97.36	98.24
	$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma$	98.9	96.1	98.67	99.34	98.91	99.42
	$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma$	97.18	98.56	99.93	99.36	99.03	98.47

Table 5.5. The ARL performance of control charts under shift in $\beta_0 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$

Chart type	**Shift type**	**Shift size, δ**					
		0.5	**1**	**1.5**	**2**	**2.5**	**3**
Shewhart	$\beta_0 + \delta\sigma, \beta_2 + \delta\sigma$	78.24	86.84	80.38	89.28	89.28	88.28
	$\beta_0 - \delta\sigma, \beta_2 - \delta\sigma$	80.14	78.68	74.73	77.08	71.88	74.19
	$\beta_0 + \delta\sigma, \beta_2 - \delta\sigma$	81.01	82.81	82.89	82.23	85.63	88.06
	$\beta_0 - \delta\sigma, \beta_2 + \delta\sigma$	80.12	82.84	74.12	79.68	80.1	77.96
CUSUM	$\beta_0 + \delta\sigma, \beta_2 + \delta\sigma$	80.12	82.84	74.12	79.68	80.1	77.96
	$\beta_0 - \delta\sigma, \beta_2 - \delta\sigma$	56.5	52.9	24.2	28.4	20	28.53
	$\beta_0 + \delta\sigma, \beta_2 - \delta\sigma$	28.6	22.5	35.2	27.6	21.6	17.42
	$\beta_0 - \delta\sigma, \beta_2 + \delta\sigma$	22.1	39	11.1	9.16	5.54	3.79
EWMA	$\beta_0 + \delta\sigma, \beta_2 + \delta\sigma$	97.21	99.28	98.82	98.69	98.61	97.28
	$\beta_0 - \delta\sigma, \beta_2 - \delta\sigma$	98.9	99.46	98.87	98.84	97.87	96.46
	$\beta_0 + \delta\sigma, \beta_2 - \delta\sigma$	99.98	99.74	99.66	99.28	99.14	98.58
	$\beta_0 - \delta\sigma, \beta_2 + \delta\sigma$	99.21	98.75	97.84	96.45	96.39	95.52

As a result, CUSUM control chart gets the smallest ARL values when two regression coefficients are shifted simultaneously especially when shift has occurred at β_1 and β_2. CUSUM chart is good at earlier detection of out-of-control observations in this case. The EWMA control chart shows the worst performance with highest ARL values.

Table 5.6. The ARL performance of control charts under shift in $\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$

Chart type	**Shift type**	**Shift size, δ**					
		0.5	**1**	**1.5**	**2**	**2.5**	**3**
Shewhart	$\beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	82.87	76.98	84.19	83.44	79.96	83.62
	$\beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	80.39	74.91	78.95	83.55	78.28	79.37
	$\beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	82.85	79.76	79.33	82.49	80.81	82.81
	$\beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	80.79	83.04	84.85	81.66	83.34	83.41
CUSUM	$\beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	36.9	21.5	11.1	8.78	2.9	2.31
	$\beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	57.3	25.6	48.8	61.7	19.2	6.12
	$\beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	34.6	30.2	9.1	2.54	3.2	3.46
	$\beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	38.6	16.89	19.5	8.2	7.4	9.01
EWMA	$\beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	99.53	99.5	99.35	98.61	98.19	96.62
	$\beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	98.38	98.27	97.72	96.83	96.16	95.81
	$\beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	99.63	99.19	98.71	98.48	97.93	97.59
	$\beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	99.12	98.91	98.06	97.85	97.79	97.72

The summary of simulation results for Shewhart, CUSUM and EWMA control charts when shifts are in all regression coefficients at the same time in different combinations are shown in Table 5.7., 5.8., and 5.9., respectively.

The key conclusion obtained from Table 5.7., 5.8., and 5.9., are as follows:

- The CUSUM control chart generally shows a good performance when all regression coefficients are shifted positively.

- Shewhart control chart did not show noticeable differences depend on shift size.
- The ARL values of EWMA control chart shows a good performance in small shift sizes and when parameters are shifted to $\beta_0 - \delta\sigma, \beta_1 + \delta\sigma, \beta_2 + \delta\sigma$ and $\beta_0 - \delta\sigma, \beta_1 - \delta\sigma, \beta_2 + \delta\sigma$.

Table 5.7. The ARL performance of Shewhart control chart under shift in $\beta_0 \pm \delta\sigma$, $\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$

Shift type	**Shift size, δ**					
	0.5	**1**	**1.5**	**2**	**2.5**	**3**
$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	75.37	86.59	84.36	85.81	80.04	87.55
$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	80.68	82.73	77.44	77.58	82.43	86.49
$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	78.02	84.25	84.31	77.87	82.05	87.02
$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	72.43	72.86	71.68	77.81	73.69	75.76
$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	84.44	85.14	78.9	78.48	83.14	84.78
$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	81.15	76.99	81.67	84.24	79.95	86.93
$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	82.37	84.3	84.66	84.68	81.72	81.96
$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	76.18	80.85	84.24	80.85	80.14	80.94

Table 5.8. The ARL performance of CUSUM control chart under shift in $\beta_0 \pm \delta\sigma$, $\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$

Shift type	**Shift size, δ**					
	0.5	**1**	**1.5**	**2**	**2.5**	**3**
$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	24.6	7.7	5.8	4.5	2.7	2
$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	43.4	26.7	42	30.2	31.3	75
$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	49.6	22.6	17.6	13.2	8.8	5
$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	29.9	13.2	4	3.1	2.2	1
$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	37.2	21.5	16.5	18.2	9.7	10
$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	45.8	33	12.9	10.7	6.2	3
$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	50.7	48.6	36.4	24.8	28.2	5
$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	35	66.2	52.1	57.9	26.6	9

Table 5.9. The ARL performance of EWMA control chart under shift in $\beta_0 \pm \delta\sigma$, $\beta_1 \pm \delta\sigma$ and $\beta_2 \pm \delta\sigma$

Shift type	**Shift size, δ**					
	0.5	**1**	**1.5**	**2**	**2.5**	**3**
$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	69.54	75.9	84.24	86.02	81.59	88.96
$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	55.93	71.34	65.46	76.42	68.36	73.48
$\beta_0 + \delta\sigma, \beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	54.36	61.86	75.44	70.38	73.84	82.75
$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma, \beta_2 + \delta\sigma$	67.17	67.96	14.65	7.01	67.96	14.65
$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma, \beta_2 + \delta\sigma$	68.7	53.93	20.16	9.12	53.93	20.16
$\beta_0 - \delta\sigma, \beta_1 + \delta\sigma, \beta_2 - \delta\sigma$	72.51	78.41	25.1	12.44	78.41	25.1
$\beta_0 + \delta\sigma, \beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	77.14	72.38	61.87	14.76	92.38	61.87
$\beta_0 - \delta\sigma, \beta_1 - \delta\sigma, \beta_2 - \delta\sigma$	82.2	67.96	14.65	7.01	67.96	14.65

5.1.1.2. Changes in the Mean

Additive shift or shifts in the mean is another type of shift which is considered in simulation out-of-control observations were obtained through changes in the mean of the generated response variables which is the inverse of link function. Computed deviance residuals from out-of-control data are used in constructing control charts.

Table 5.10. The *ARL* performance of control charts under shift in $\lambda \pm \delta\sigma$

Chart type	**Shift type**	**Shift size, δ**					
		0.5	**1**	**1.5**	**2**	**2.5**	**3**
Shewhart	$\lambda + \delta\sigma$	63.79	49.53	34.65	28.6	21.99	23.91
	$\lambda - \delta\sigma$	86.94	85.29	87.46	78.83	88.67	88.53
CUSUM	$\lambda + \delta\sigma$	32.6	33.9	22.1	19.6	16.5	3.3
	$\lambda - \delta\sigma$	29.5	41.6	23.4	17.2	10.06	3.12
EWMA	$\lambda + \delta\sigma$	40.26	16.36	7.66	5.4	3.53	3.77
	$\lambda - \delta\sigma$	99.02	96.08	93.13	95.1	98.03	97.06

The simulation results are summarized in Table 5.10. The conclusions obtained are as follows:

- Shift values up to 1.5σ, the Shewhart control chart shows a good performance when the mean is shifted positively.
- The CUSUM control chart shows a good performance when the mean is shifted negatively.
- Despite the good performance when the mean is shifted positively, the EWMA control chart shows the worst result when the mean is shifted negatively.

5.2. Real Case Study

In this section, the monitoring method of count data is applied to a timber industry which manufactures laminated plastic plywood. The study consisted of evaluating the effect of input variables over the number of defects found in produced plywood. The quality of plywood is related to some variables, as detailed by Dermirkir et al. (2013), Fang et al. (2014) and Azaman et al. (2013). The response variable is the number of defects per laminated plastic plywood area (y) and the predicted variables are volumetric shrinkage (x_1), assembly time (x_2), wood density (x_3), and drying temperature (x_4).

Figure 5.1. shows the result of Kolmogorov-Smirnov test which indicates that the response variable follow the Poisson distribution.

One-Sample Kolmogorov-Smirnov Test

		response
N		100
Poisson Parameter[a,b]	Mean	14.2800
Most Extreme Differences	Absolute	.658
	Positive	.658
	Negative	-.118
Kolmogorov-Smirnov Z		6.580
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Poisson.

b. Calculated from data.

Figure 5.1. Kolmogorov-Smirnov test result of response variables

Table 5.11. contains the deviance residuals which are obtained from data and used in the construction of control charts.

The $k = 2.6$ for Shewhart control chart, the $h = 2.59$ for CUSUM control chart and the $\lambda = 0.7$ and $A^r = 3.54$ for EWMA control chart are determined so as to achieve ARL_0=100. The lower and upper control limits for Shewhart control chart are obtained as -2.6883 and 2.5809, for CUSUM control chart are obtained as -2.6245 and 2.6245 and for EWMA control chart are obtained as -2.6833 and 2.5759.

Control charts are shown in Figure 5.2., 5.3., and 5.4 where Figure 5.2. shows the Shewhart control chart, Figure 5.3. shows the CUSUM control chart and Figure 5.4. shows the EWMA control chart.

Table 5.11. Deviance residuals obtained from data

N	Deviance residual	N	Deviance residual	N	Deviance residual	N	Deviance residual	N	Deviance residual
1	-0.12759	21	-2.58865	41	0.466695	61	-1.41262	81	0.496411
2	0.646995	22	-0.21348	42	0.17173	62	-1.57482	82	0.056015
3	0.992477	23	-1.22396	43	0.709548	63	-0.51907	83	0.195433
4	-0.49535	24	-0.25379	44	-0.20816	64	-1.17923	84	0.134732
5	1.611807	25	-0.04937	45	-2.06343	65	-0.29489	85	0.194238
6	0.455879	26	0.257111	46	0.371368	66	-0.72256	86	0.422045
7	1.570317	27	0.024604	47	1.44307	67	2.679134	87	1.048551
8	-2.08736	28	-2.04392	48	2.195156	68	-0.71689	88	-0.0231
9	-0.57927	29	-1.07767	49	-1.42777	69	0.011384	89	-0.10746
10	-0.11085	30	0.337518	50	0.639222	70	-0.22887	90	-0.87062
11	-0.65025	31	0.460817	51	-0.23402	71	0.799774	91	-1.0121
12	-1.74044	32	-1.1093	52	0.24814	72	-1.292	92	-1.0298
13	0.189972	33	-0.03083	53	0.78237	73	-1.63914	93	0.147633
14	0.426147	34	0.634027	54	-0.17854	74	-1.79486	94	-0.38305
15	0.923271	35	0.744937	55	-0.2971	75	0.433483	95	1.541056
16	0.031388	36	1.004069	56	0.963125	76	-0.57881	96	-0.16081
17	0.353547	37	-0.7132	57	1.269221	77	1.202645	97	2.253991
18	-1.46867	38	-0.78323	58	0.768824	78	0.954593	98	0.064904
19	0.023225	39	-0.06761	59	-1.2475	79	-0.8952	99	0.188231
20	-1.28537	40	2.032338	60	0.608162	80	-0.20801	100	-0.55281

The Figure 5.3. shows that the CUSUM control chart is much more sensitive than others and quickly detects out-of-control observations.

The Shewhart control chart is also detected the out-of-control observations whereas, the EWMA control chart shows that the control chart statistics are not exceeded control limits, in other words, all observations are in control.

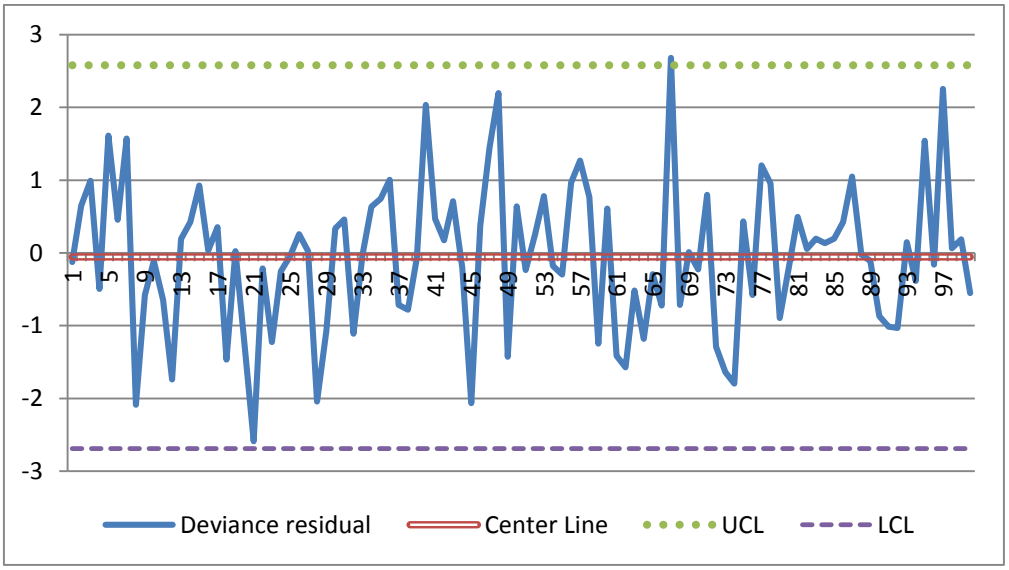


Figure 5.2. Shewhart control chart

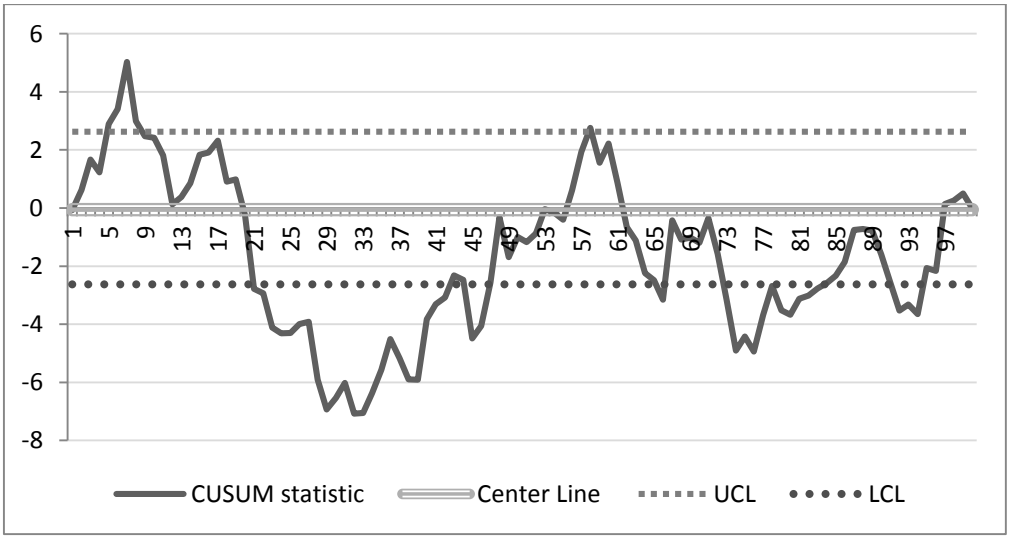


Figure 5.3. CUSUM control chart

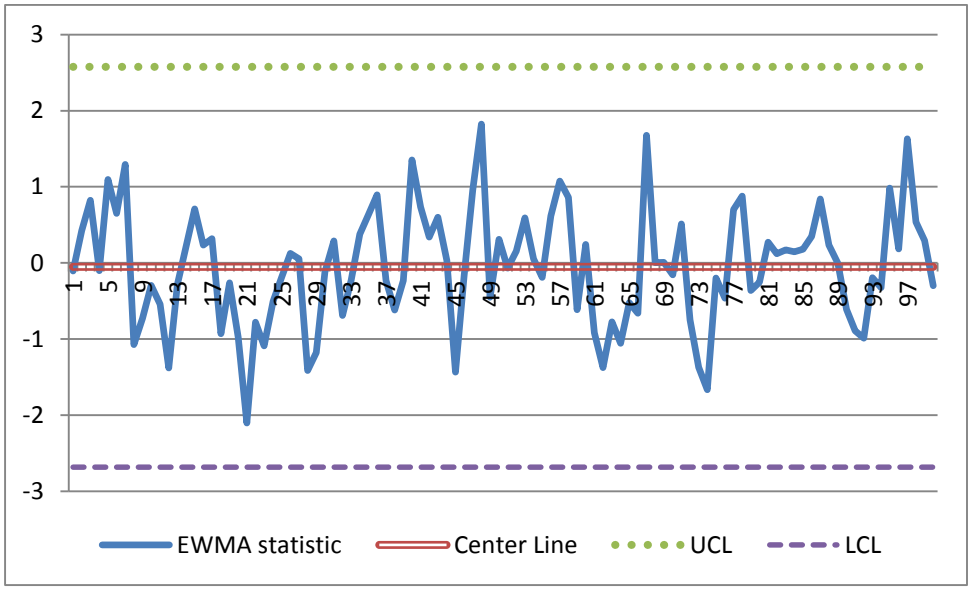


Figure 5.4. EWMA control chart

6. CONCLUSION

Control charts are the most used tools of statistical process control and a fundamental assumption of traditional control charts is that the observations of the process are independent and identically normal distributed about a certain mean. Because the assumptions are not realistic in practice, during the many years researchers have proposed some modifications for traditional control charts. One of these propositions is using regression for creating control charts. A control chart usage of regression model is a quality tool for monitoring response variables that vary as a result of modifications in the control variables. The tool consists of building a regression model and obtaining the control limits using the deviance residuals from the model. GLM based control charts are used when response variables do not follow the normal distribution.

The aim of this thesis is to compare GLM based control charts which have been used for monitoring of the count data.

In this thesis, firstly, an overview of GLM and the information of Poisson regression model are given. Then, the traditional control charts are discussed and an extinction of these control charts to GLM especially when response variables are Poisson distributed are examined.

Performances of Poisson Shewhart, Poisson CUSUM and Poisson EWMA control charts are compared by using average run length performance criterion through the simulations. It can be concluded that none of the compared charts has an overall performance. Every single control chart has good performances for some situations, and bad performances for some.

Shewhart control chart quickly detects out-of-control observations when β_2 is shifted. In high shift values, Shewhart and EWMA control chart shows a good performance when mean is shifted positively. The CUSUM control chart outperformed when all regression coefficients are shifted positively. This

control chart also shows a good performance when the mean is shifted negatively.

When mean is shifted positively the Shewhart control chart shows a good performance in high shift sizes. The CUSUM control chart shows a good performance when the mean is shifted negatively.

As a result, changes in ARL value shows the CUSUM control chart is much more sensitive to multiplicative shifts, unlike Shewhart and EWMA control charts. Shewhart control chart gets high ARL values despite the shifts in any situation and shift size shows that the Shewhart and EWMA control charts are not good choice for the detection of the out-of-control observations when response variables follow the Poisson distribution.

The overall result indicates that the Poisson based CUSUM control chart has good accuracy in detecting changes in count response. However, in general, the CUSUM control chart has best performance in most of the shift combinations. The real case study also showed that the CUSUM control chart is good at quick detecting out of control observations.

REFERENCES

Agresti, A., Amiri, A., and Niaki, S.T.A., 2014. A New Link Function in Generalized Linear Model Based Control Charts to Improve Monitoring of Two-Stage Processes with Poisson Response. International Journal of Advanced Manufacturing Technology, 72:1243-1256.

Agresti, A., 2015. Foundations of Linear and Generalized Linear Models. John Wiley and Sons, Hoboken, 472pgs.

Altham, P.M.E., 2011. Introduction to Generalized Linear Modelling. Statistical Laboratory, University of Cambridge, 104pgs.

Amiri, A., Koosha, M., Azhdari, A., and Wang, G., 2015. Phase I Monitoring of Generalized Linear Model Based Regression Profiles. Journal of Statistical Computation and Simulation, 85(14):2839-2859.

Asadzadeh, S., Yang, S., and Aghaie, A., 2008. Monitoring and Diagnosing Multistage Processes: A Review of Cause Selecting Control Charts. International Journal of Industrial and Systems Engineering, 2(3):214-235.

Azaman, M.D., Sapuan, S.M., Sulaiman, S., Zainudin, E.S., and Khalina, A., 2013. Shrinkages and Warpage in the Processability of Wood-filled Polypropylene Composite Thin-walled Parts Formed by Injection Molding. Materials & Design, 52:1018–1026.

Barnard, G.A., 1959. Control Charts and Stochastic Processes. Journal of Royal Statistical Society, 21:239–271.

Box, G.E.P., Jenkins, G.M., and Macgregor, J.F., 1974. Some Recent Advances in Forecasting and Control. Applied Statistics, 23: 158-179.

Benbow, D.W., and Broome, H.W., 2013. The Certified Reliability Engineer Handbook, 2nd Edition. Quality Press, 360pgs.

Brook, D., and Evans, D.A., 1972. An Approach to the Probability Distribution of CUSUM Run Length. Biometrika, 59:539-549.

Cameron, A.C., and Trivedi, P.K., 2013. Regression Analysis of Count Data, 2nd Edition. Cambridge University Press, 588pgs.

Cisar, P., and Cisar, S.M., 2011. Optimization Methods of EWMA Statistics. Acta Polytechnica Hungarica. 8(5):73-87.

Crosier, R.B., 1988. Multivariate Generalizations for Cumulative Sum Quality-Control Schemes. Technometrics, 30:291–303.

Crowder, S.V., 1987. A Simple Method for Studying Run Length Distributions of Exponentially Weighted Moving Average Charts. Technometrics, 29:401-407.

_____, 1989. Design of Exponentially Weighted Moving Average Schemes. Journal of Quality Technology, 21:155–162.

Demirkir, C., Ozsahin, S., Aydin, I., and Colakoglu, G., 2013. Optimization of Some Panel Manufacturing Parameters for the Best Bonding Strength of Plywood. International Journal of Adhesion & Adhesives, 46:14–20.

DiPaola, P.P., 1945. Use of Correlation in Quality Control. Industrial Quality Control. 2(1):10-14.

Dobson, A.J., and Barnett, A.G., 2008. An Introduction to Generalized Linear Models, 3rd Edition. Chapman & Hall/ CRC Texts in Statistical Science.

Fang, L., Chang, L., Guo, W.J., Chen, Y., and Wang, Z., 2014. Influence of Silane Surface Modification of Veneer on Interfacial Adhesion of Wood–Plastic Plywood. Applied Surface Science, 288:682–689.

Ferguson, D.L., 2013. Removing the Barriers to Efficient Manufacturing: Real-World Applications of Lean Productivity. Productivity Press, 294pgs.

Filho, D.M., and Anna, A.M.O.S., 2015. Principal Components Regression Based Control Charts for Monitoring Count Data. International Journal of Advanced Manufacturing Technology, 85:1565-1574.

Hauck, D.J., Runger, G.C., and Montgomery, D.C., 1999. Multivariate Statistical Process Monitoring and Diagnosis with Grouped Regression-Adjusted

Variables. Communications in Statistics, Simulation and Computation, 28(2):309-328.

Hawkins, D.M., 1991. Multivariate Quality Control Using Regression Adjusted Variables. Technometrics, 33:61-75.

Hawkins, D.M., 1993. Regression Adjustment for Variables in Multivariate Quality Control. Journal of Quality Technology, 25(3):170-182.

Haworth, D.A., 1996. Regression Control Charts to Manage Software Maintenance. Software Maintenance: Research and Practice, 8:35-48.

_____, 1993. Regression Adjustment for Variables in Multivariate Quality Control. Journal of Quality Technology, 25:170-182.

Hoblos, J., 2006. Selfish Node Misbehaving Statistical Detectıon with Active Mac Layer Nav Attack in Wireless Networks. Kenn State University.

Hotelling, H., 1947. Multivariate Quality Control - Techniques of Statistical Analysis (edited by Eisenhart, C., Hastay, M.W., and Wallis, W.A., McGraw-Hill). New York, 113–184.

Hunter, J. S., 1986. The Exponentially Weighted Moving Average. Journal of Quality Technology, 18: 203-210.

Jackson, J.E., 1956. Quality control methods. Industrial Quality Control, 4-8.

Jearkpaporn, D., Montgomery, D.C., Runger, G.C., and Borror C.M., 2003. Process Monitoring for Correlated Gamma Distributed Data Using Generalized Linear Model Based Control Charts. Quality and Reliability Engineering International, 19:477-491.

_____, 2005. Model Based Process Monitoring Using Robust Generalized Linear Models. International Journal of Production Research, 43(7):1337-1354

_____, 2007. Process Monitoring for Mean Shifts for Multiple Stage Processes. International Journal of Production Research, 45(23):5547-5570.

Kleinbaum, D., Kupper, L., Nizam, A., and Muller, K., 1998. Applied Regression Analysis and Other Multivariable Methods, 3rd Edition. Duxbury Press.

Li, Z., Zou, C., Gong, Z., and Wang, Z., 2014. The Computation of Average Run Length and Average Time to Signal: An Overview. Journal of Statistical Computation and Simulation, 84(8): 1779-1802.

Lindsey, J.K., 1997. Applying Generalized Linear Models. Springer. New York, 256pgs.

Loredo, E.N., Jearkpaporn, D., and Borror, C.M., 2002. Model Based Control Chart for Autoregressive and Correlated Data. Quality and Reliability Engineering International, 18:489-496.

Lowry, C.A., Woodall, W.H., Champ, C.W., and Rigdon, S.E., 1992. A Multivariate Exponentially Weighted Moving Average Control Chart. Technometrics, 34:46-53.

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

Lucas, M.J., 1976. The Design and Use of V-Mask Control Schemes. Quality Technology, 14:51-59.

_____, 1985. Counted Data CUSUMs. Technometrics, 27:129-144.

Mandel, B.J., 1969. The Regression Control Chart. Journal of Quality Technology, 1:1-9.

May, J.G., and Sulek, J.M., 2008. Regression Control Charts for Limited Process Data. 2008 Southeast Decision Sciences Institute Conference Proceedings, 667-677.

McCullagh, P., and Nelder, J.A., 1989. Generalized Linear Models, 2nd Edition. Chapman and Hall, 532pgs.

Montgomery, D.C., Peck, E.A., and Vining, G., 2012. Introduction to Linear Regression Analysis, 5th Edition. John Wiley & Sons, 672pgs.

Montgomery, D.C., 2009. Introduction to Statistical Quality Control, 6th Edition. John Wiley & Sons, 734pgs.

Myers, R.H., 2000. Classical and Modern Regression with Applications, 2nd Edition. Duxbury Classic, 488pgs.

Myers, R.H., Montgomery, D.C., Vining, G., and Robinson, T.J., 2010. Generalized Linear Models: with Applications in Engineering and the Sciences, 2nd Edition. John Wiley & Sons, 496pgs.

Nelder, J.A., and Wedderburn, R.W.M., 1972. Generalized Linear Models. Journal of the Royal Statistical Society, 135(3):370-384.

NIST/SEMATECH, 2012. E-Handbook of Statistical Methods. Published online, URL: http://www.itl.nist.gov/div898/handbook/

Page, E.S., 1954. Continuous Inspection Schemes. Biometrika, 41(1):100–115.

Pehlivan, C., 2008. Controlling High-Quality Manufacturing Processes: A Robustness Study of the Lower-Sided TBE EWMA Procedure. METU.

Pierce, D., and Schafer, D., 1986. Residuals in Generalized Linear Models. Journal of the American Statistical Association, 81(396): 977–986.

Rao, C.R., 1947. Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Applications to Problems of Estimation. Proceedings of the Cambridge Philosophical Society, 44:50–57.

Roberts, S.W., 1959. Control Chart Tests Based on Geometric Moving Averages. Technometrics, 1:239-250.

_____, 1966. A Comparison of Some Control Chart Procedures. Technometrics, 8:411-430.

Robinson, P. B., and Ho, T.Y., 1978. Average Run Lengths of Geometric Moving Average Charts by Numerical Methods. Technometrics, 20:85-93.

Rodriguez, G., 2007. Lecture Notes on Generalized Linear Models. Published online. URL: http://data.princeton.edu/wws509/notes/

Ryan, T.P., 2000. Statistical Methods for Quality Improvement, 2nd Edition. John Wiley & Sons, New York, 592pgs.

Santos-Fernandez, E., 2013. Multivariate Statistical Quality Control Using R. Springer, New York, 127pgs.

SAS Institute Inc., 2012. JMP 10 Quality and Reliability Methods. SAS Institute Inc., 423pgs.

Sauza, F.S., Caten, C. S., and Pendirini, D., 2015. Capability Indices for Control Charts Based on Regression Models. Brazilian Journal of Operations and Management. 12:234-247.

Shadman, A., Mahlooji, H., Yeh, A.B., and Zou C., 2015. A Change Point Method for Monitoring Generalized Linear Profiles in Phase I. Quality and Reliability Engineering International, 31:1367-1381.

Shu, L., and Tsung, F., 2003. On Multistage Statistical Process Control. Journal of the Chinese Institute of Industrial Engineers, 20(1):1-8.

Shu, L., Tsui, K.L., and Tsung, F., 2004. Run Length Perfomance of Regression Control Charts with Estimated Parameters. Journal of Quality Technology, 36:280-292.

_____, 2005. Effect of Estimation Errors on Cause Selecting Charts. IIE Transactions, 37(6):559-567.

_____, 2008. A Review of Regression Control Charts. Encyclopedia of Statistics in Quality and Reliability, 260:1–9.

Skinner, K.R., Montgomery, D.C., and Runger, G.C., 2003. Process Monitoring for Multiple Count Data Using Generalized Linear Model Based Control Charts. International Journal of Production Research, 41(6):1167-180.

_____, 2004. Generalized Linear Model Based Control Charts for Discrete Semiconductor Process Data. Quality and Reliability Engineering International, 20(8): 777- 786.

_____, 2006. Process Monitoring for Multiple Count Data Using a Deleted-Y Statistic. Quality Technology & Quantitative Management, 3(3): 247-262.

Smyth, G. K., 2003. Pearson's Goodness of Fit Statistic as A Score Test Statistic. Science and Statistics: A Festschrift for Terry Speed (edited by Goldstein, D. R.). Institute of Mathematical Statistics, Ohio, 40:115-126.

Stoumbos, Z.G., Reynolds, M.R., and Woodall, W.H., 2003. Control Chart Schemes for Monitoring the Mean and Variance of Processes. Handbook of Statistics: Statistics in Industry. Elsevier Science, Amsterdam, 553-571.

Suhasini, S.R., 2006. Lecture notes in Generalized Linear Models. Published online.

Sulek, J.M., Marucheck, A., and Lind, M.R., 2006. Measuring Performance in Multistage Service Operations: An Application of Cause Selecting Control Charts. Journal of Operations Management, 24(5):711-727.

Testik, M.C., and Runger, G.C., 2006. Multivariate One-sided Control Charts. IIE Transactions, 38(8):635-645.

Wade, M.R., and Woodall, W.H., 1993. A Review and Analysis of Cause Selecting Control Chart. Journal of Quality Technology, 25(3):161-169.

Waldmann, K.H., 1986. Bounds for the Distribution of the Run Length of Geometric Moving Average Charts. Applied Statistics - Journal of the Royal Statistical Society Series C, 35: 151-158

Weis, P.E., 1957. An Application of a Two-Way X-Bar Chart. Industrial Quality Control, 14(6):23-37.

Woodall, W.H., and Ncube, M.M., 1985. Multivariate CUSUM Quality Control Procedures. Technometrics, 27:285-292.

Woodall, W.H., 1997. Control Charting Based on Attribute Data: Bibliography and Review. Journal of Quality Technology, 29:172-183.

Yang S.F., and Yang, C.M., 2005. Effects of Imprecise Measurement on the Two Dependent Processes Control for the Autocorrelated Observations. International Journal of Advanced Manufacturing Technology, 26(5-6):623-630.

_____, 2006. An Approach to Controlling Two Dependent Process Steps with Autocorrelated Observations. International Journal of Advanced Manufacturing Technology, 29(1-2):170-177.

Zhang, G.X., 1980. A New Type of Quality Control Chart Allowing the Presence of Assignable Causes the Cause Selecting Control Chart. ACTA ELECTRON-ICA SANICA, 2:1-10.

_____, 1984. A New Type of Control Charts and Theory of Diagnosis with Control Charts. World Quality Congress Transactions, American Society for Quality Control, 175-185.

Zeng, L., and Zhou, S., 2008. Variability Monitoring of Multistage Manufacturing Processes Using Regression Adjustment Methods. IIE Transactions, 40(2):109-121.

http://isoconsultantpune.com/statistical-process-control-using-control-chart/

CURRICULUM VITAE

Ulduz MAMMADOVA was born in Baku/Azerbaijan on August 4, 1992. She graduated from Physics - mathematics and Informatics weighted high school in 2006. Then, she started her Bachelor Degree in 2009 at Azerbaijan State Economic University, Faculty of Economics and Accountancy, Department of Statistics. In 2014 she started her MSc education at Cukurova University, Department of Statistics.

APPENDICES

Appendix 1

Constants for Control Charts (Benbow et al., 2013)

Sample size	Constant values												
N	**A2**	**A3**	**c4**	**B3**	**B4**	**B5**	**B6**	**d2**	**d3**	**D1**	**D2**	**D3**	**D4**
2	1.880	2.659	0.7979	0.000	3.267	0.000	2.606	1.128	0.853	0.000	3.686	0.000	3.267
3	1.023	1.954	0.8862	0.000	2.568	0.000	2.276	1.693	0.888	0.000	4.358	0.000	2.575
4	0.729	1.628	0.9213	0.000	2.266	0.000	2.088	2.059	0.880	0.000	4.698	0.000	2.282
5	0.577	1.427	0.9400	0.000	2.089	0.000	1.964	2.326	0.864	0.000	4.918	0.000	2.114
6	0.483	1.287	0.9515	0.030	1.970	0.029	1.874	2.534	0.848	0.000	5.079	0.000	2.004
7	0.419	1.182	0.9594	0.118	1.882	0.113	1.806	2.704	0.833	0.205	5.204	0.076	1.924
8	0.373	1.099	0.9650	0.185	1.815	0.179	1.751	2.847	0.820	0.388	5.307	0.136	1.864
9	0.337	1.032	0.9693	0.239	1.761	0.232	1.707	2.970	0.808	0.547	5.394	0.184	1.816
10	0.308	0.975	0.9727	0.284	1.716	0.276	1.669	3.078	0.797	0.686	5.469	0.223	1.777
11	0.285	0.927	0.9754	0.321	1.679	0.313	1.637	3.173	0.787	0.811	5.535	0.256	1.744
12	0.266	0.886	0.9776	0.354	1.646	0.346	1.610	3.258	0.778	0.923	5.594	0.283	1.717
13	0.249	0.850	0.9794	0.382	1.618	0.374	1.585	3.336	0.770	1.025	5.647	0.307	1.693
14	0.235	0.817	0.9810	0.406	1.594	0.399	1.563	3.407	0.763	1.118	5.696	0.328	1.672
15	0.223	0.789	0.9823	0.428	1.572	0.421	1.544	3.472	0.756	1.203	5.740	0.347	1.653
16	0.212	0.763	0.9835	0.448	1.552	0.440	1.526	3.532	0.750	1.282	5.782	0.363	1.637
17	0.203	0.739	0.9845	0.466	1.534	0.458	1.511	3.588	0.744	1.356	5.820	0.378	1.622
18	0.194	0.718	0.9854	0.482	1.518	0.475	1.496	3.640	0.739	1.424	5.856	0.391	1.609
19	0.187	0.698	0.9862	0.497	1.503	0.490	1.483	3.689	0.733	1.489	5.889	0.404	1.596
20	0.180	0.680	0.9869	0.510	1.490	0.504	1.470	3.735	0.729	1.549	5.921	0.415	1.585
21	0.173	0.663	0.9876	0.523	1.477	0.516	1.459	3.778	0.724	1.606	5.951	0.425	1.575
22	0.167	0.647	0.9882	0.534	1.466	0.528	1.448	3.819	0.720	1.660	5.979	0.435	1.565
23	0.162	0.633	0.9887	0.545	1.455	0.539	1.438	3.858	0.716	1.711	6.006	0.443	1.557
24	0.157	0.619	0.9892	0.555	1.445	0.549	1.429	3.895	0.712	1.759	6.032	0.452	1.548
25	0.153	0.606	0.9896	0.565	1.435	0.559	1.420	3.931	0.708	1.805	6.056	0.459	1.541

APPENDIX 2

Reference samples of data used in plastic plywood case study (Filho et al., 2015)

y - the number of defects per laminated plastic plywood area; x_1 - volumetric shrinkage; x_2 - assembly time; x_3 - wood density; x_4 - drying temperature

N	y	x_1	x_2	x_3	x_4	N	y	x_1	x_2	x_3	x_4
1	18	12.39	17.43	0.52	117.62	**26**	17	10.27	15.52	0.55	136.05
2	1	8.49	14.15	0.52	106.45	**27**	1	10.02	15.64	0.51	97.45
3	41	9.83	15.13	0.57	149.61	**28**	6	10.66	15.82	0.54	129.94
4	53	12.39	17.9	0.54	123.93	**29**	13	10.42	15.48	0.55	137.11
5	3	8.58	14.03	0.53	117.65	**30**	3	8.16	13.79	0.55	131.98
6	24	10.66	16.17	0.55	131.58	**31**	4	7.78	13.36	0.56	139.81
7	26	11.18	16.55	0.54	125.56	**32**	33	10.53	15.83	0.56	144.74
8	0	8.47	13.75	0.53	142.64	**33**	2	7.89	13.21	0.55	137.69
9	4	9.78	15.03	0.54	127.06	**34**	16	10.69	16.18	0.54	127.07
10	4	10.85	16.12	0.52	112.38	**35**	18	10.18	15.5	0.55	135.66
11	8	11.11	16.31	0.53	121.27	**36**	1	7.9	13.48	0.52	108.93
12	0	9.18	14.88	0.53	111.07	**37**	6	10.23	15.34	0.54	128.9
13	12	9.77	15	0.55	139.21	**38**	1	10.06	15.36	0.52	111.13
14	30	10.09	15.47	0.56	144.64	**39**	1	8.56	14.16	0.53	117.9
15	2	7.77	13.3	0.54	126.4	**40**	5	7.74	13.45	0.55	131.5
16	5	9.74	15.19	0.54	123.61	**41**	8	10.06	15.38	0.54	125.97
17	15	10.03	15.46	0.55	135.28	**42**	9	10.9	16.23	0.53	120.29
18	28	11.16	16.44	0.55	136.21	**43**	2	10.03	14.95	0.51	110.89
19	1	8.91	14.02	0.52	122.04	**44**	4	11.06	16.5	0.52	106.44
20	0	8.58	14.04	0.53	112.62	**45**	26	10.52	15.76	0.56	144.57
21	6	10.06	15.38	0.54	146.16	**46**	404	12.28	17.53	0.57	155.71
22	1	8.7	14.07	0.53	121.27	**47**	7	10.22	15.78	0.53	112.65
23	0	7.51	13.36	0.54	121.79	**48**	2	7.47	12.62	0.52	119.97
24	9	10.39	15.85	0.54	125.98	**49**	0	9.32	14.73	0.52	109.57
25	4	10.99	16.22	0.52	108.85	**50**	2	8.59	13.91	0.53	124.46

N	y	x_1	x_2	x_3	x_4	N	y	x_1	x_2	x_3	x_4
51	87	12.08	17.17	0.55	140.29	**76**	4	10.28	15.6	0.53	122.6
52	6	11.11	16.27	0.52	114.15	**77**	2	9.15	14.98	0.52	99.82
53	4	7.97	13.4	0.55	140.9	**78**	5	8.48	14.08	0.55	131.42
54	1	8.85	14.11	0.53	118.37	**79**	1	11.78	17.24	0.5	87.34
55	4	9.49	14.62	0.54	132.04	**80**	2	10.12	15.42	0.52	113.44
56	2	9.41	15.06	0.52	101.86	**81**	1	8.34	14.33	0.53	102.17
57	10	8.98	14.32	0.55	141.84	**82**	6	10.57	15.91	0.53	118.3
58	1	9.17	14.87	0.51	94.14	**83**	1	8.5	14.12	0.53	112.09
59	1	9.67	15.36	0.53	115.73	**84**	15	10.73	15.97	0.54	131.08
60	6	9.63	14.98	0.54	125.89	**85**	3	8.28	13.72	0.55	133.94
61	1	9.16	14.34	0.54	129.12	**86**	1	8.56	13.71	0.52	118.57
62	15	9.21	14.36	0.57	154.07	**87**	6	8.65	14.29	0.55	131.59
63	1	7.46	12.63	0.55	143.61	**88**	9	9.79	15.21	0.55	130.42
64	0	8.04	13.46	0.53	121.33	**89**	2	8.03	13.4	0.55	135.13
65	9	8.87	14.19	0.56	147.4	**90**	0	9.04	14.34	0.51	100.18
66	0	8.74	14.84	0.51	86.97	**91**	0	9.89	15.34	0.5	94.26
67	29	10.75	15.85	0.54	136.39	**92**	3	10.5	15.88	0.53	116.07
68	4	8.57	14.11	0.56	136.79	**93**	115	12.19	16.93	0.55	148.56
69	1	9.42	14.88	0.52	105.78	**94**	4	10.23	15.72	0.53	119.18
70	20	10.76	16.28	0.55	128.15	**95**	15	10.31	15.79	0.54	127.62
71	118	12.18	16.96	0.55	147.36	**96**	6	7.87	13.34	0.57	149.34
72	1	9.82	15.21	0.53	116.98	**97**	10	9.32	14.19	0.54	138.5
73	4	9.46	14.62	0.55	139.07	**98**	1	8.35	13.66	0.53	123.23
74	0	8.37	13.42	0.54	132.05	**99**	1	8.52	14.06	0.53	112.88
75	2	9.35	14.31	0.52	123.9	**100**	2	10.4	15.59	0.52	114.12

APPENDIX 3

MATLAB codes

```
%%%%%% Performance Analysis of Shewhart control chart

clc;clear;
%%%constants
n=100; m=2; L=0;
NPrelim=1000; Nsim=1000;
RL=[];

%% input variables
X=randn(n,m);
X1=X(:,1); X2=X(:,2);

%% coefficients in-control
beta0=1; beta1=2; beta2=1;
Beta=[beta0;beta1;beta2];

%% mean of the poisson distribution/inverse link function
lambda=exp(beta0+beta1*X1+beta2*X2);

%% simulation loop:in-control mean and standard deviation are
simulated based on in-control preliminary samples.
M=0; Sigma=0;

for i=1:NPrelim
    Y=poissrnd(lambda);
    mdl = fitglm(X,Y,'linear','Distribution','poisson');
    r = mdl.Residuals.deviance;
    M=M+mean(r);
    Sigma=Sigma+std(r);
end

M_in=M/NPrelim;
S_in=Sigma/NPrelim;

k=2.85;
```

```
Ucl=M_in+k*(S_in);
Cl=M_in;
Lcl=M_in-k*(S_in);

P_beta=normcdf((Ucl-M_in)/(S_in))-normcdf((Lcl-M_in)/(S_in));
ARL0=1/(1-P_beta);

shift=L*S_in;

%% Multiplicative shift - change in any regression coefficients
beta0_s=beta0+shift; beta1_s=beta1+shift; beta2_s=beta2;

%Additive shift - "mean+shift" where mean=g^-1()
%lambda_s=exp(beta0+beta1_s*X1+beta2_s*X2)+shift;

lambda_s=exp(beta0_s+beta1_s*X1+beta2_s*X2);

for i=1:Nsim
    Y2=poissrnd(lambda_s);
    mdl2 = fitglm(X,Y2,'linear','Distribution','poisson');
    r2 = mdl2.Residuals.deviance;

    for j=1:n
        if r2(j)>Ucl;
            RL(i)=j-1;
            break
        elseif r2(j)<Lcl;
               RL(i)=j-1;
            break
        else
            RL(i)=n;
        end

    end
end
     ARL=sum(RL)/Nsim

%%Control chart
%%%%%% Perfromance Analysis of EWMA control chart
```

```
clc;clear;
%%%%%%% Constants %%%%%%
n=100; m=2; L=0.25;
NPrelim=1000; Nsim=1000;
RL=[]; z=[];

%%%%%%% input variables
X=randn(n,m);
X1=X(:,1); X2=X(:,2);

%% coefficients in-control
beta0=1; beta1=2; beta2=1;
Beta=[beta0;beta1;beta2];

%% mean of the poission distribution/inverse link function
lambda=exp(beta0+beta1*X1+beta2*X2);

%% simulation loop:in-control mean and standard deviation are
simulated based on in-control preliminary samples.
M=0; Sigma=0;

for i=1:NPrelim
    Y=poissrnd(lambda);
    mdl = fitglm(X,Y,'linear','Distribution','poisson');
    r = mdl.Residuals.deviance;
    M=M+mean(r);
    Sigma=Sigma+std(r);
end

    M_in=M/NPrelim;
    S_in=Sigma/NPrelim;

    %%%%%%%%%%     Control Limits    %%%%%%%%%%%

    lambda_e=0.7;
    A_ewma =3.85*S_in;

    Ucl=M_in+A_ewma*sqrt((lambda_e/(2-lambda_e))*(S_in)^2);
    Cl=M_in;
```

```
    Lcl=M_in-A_ewma*sqrt((lambda_e/(2-lambda_e))* (S_in)^2);


    %%%%%%%%%%   In control ARL    %%%%%%%%%%

P_beta=normcdf((Ucl-M_in)/(S_in))-normcdf((Lcl-M_in)/(S_in));
ARL0=1/(1-P_beta);

    %%%%%%%%%%   Shift    %%%%

    shift=L*S_in;

%% Multiplicative shift - change in any regression coefficients
beta0_s=beta0+shift; beta1_s=beta1+shift; beta2_s=beta2;

%Additive shift - "mean+shift" where mean=g^-1()
%lambda_s=exp(beta0+beta1_s*X1+beta2_s*X2)+shift;

lambda_s=exp(beta0_s+beta1_s*X1+beta2_s*X2);

for i=1:Nsim
    Y2=poissrnd(lambda_s);
    mdl2 = fitglm(X,Y2,'linear','Distribution','poisson');
    r2 = mdl2.Residuals.deviance;
    for j=1:n
       if j<=1;
          z(j)=lambda_e*r2(j);
       elseif j>1;
          z(j)=lambda_e*r2(j)+(1-lambda_e)*z(j-1);
       end
    end
    for j=1:n
        if z(j)>Ucl;
            RL(i)=j-1;
            break
        elseif z(j)<Lcl;
            RL(i)=j-1;
            break
        else
            RL(i)=n;
        end
```

```
    end
end

ARL=sum(RL)/Nsim


%%%%%% Performance Analysis of CUSUM control chart

clc;clear;
%%%%%%% Constants %%%%%%
n=100; m=2; L=0;
NPrelim=1000; Nsim=1000;
RL=[]; Cplus=[];Cminus=[];

%%%%%%% input variables
X=randn(n,m);
X1=X(:,1); X2=X(:,2);

%% coefficients in-control
beta0=1; beta1=2; beta2=1;
Beta=[beta0;beta1;beta2];

%% mean of the poission distribution/inverse link function
lambda=exp(beta0+beta1*X1+beta2*X2);

%% simulation loop:in-control mean and standard deviation are
simulated based on in-control preliminary samples.
M=0; Sigma=0;

for i=1:NPrelim
    Y=poissrnd(lambda);
    mdl = fitglm(X,Y,'linear','Distribution','poisson');
    r = mdl.Residuals.deviance;
    M=M+mean(r);
    Sigma=Sigma+std(r);

end

    M_in=M/NPrelim;
    S_in=Sigma/NPrelim;
```

```
    k=(1/2)*S_in;

    %%%%%%%%%%     Control Limits    %%%%%%%%%%%%
    h=2.92;

    Lcl=-h*S_in;
    Ucl=+h*S_in;

    %%%%%%%%%%   In control ARL    %%%%%%%%%%

    P_beta=normcdf((Ucl-M_in)/(S_in))-normcdf((Lcl-
M_in)/(S_in));
    ARL0=1/(1-P_beta);

    %%%%%%%%%%%   Shift    %%%%

    shift=L*S_in;

%% Multiplicative shift - change in any regression coefficients
beta0_s=beta0+shift; beta1_s=beta1+shift; beta2_s=beta2;

%Additive shift - "mean+shift" where mean=g^-1()
%lambda_s=exp(beta0+beta1_s*X1+beta2_s*X2)+shift;

lambda_s=exp(beta0_s+beta1_s*X1+beta2_s*X2);

for i=1:Nsim
    Y2=poissrnd(lambda_s);
    mdl2 = fitglm(X,Y2,'linear','Distribution','poisson');
    r2 = mdl2.Residuals.deviance;

    for j=1:n
        if j<=1;
            Cplus1(j)= max(0,r21(j)-M_in-k);
            Cminus1(j)= max(0,M_in-k-r21(j));
        elseif j>1;
            Cplus1(j)= max(0,r21(j)-M_in-k+Cplus1(j-1));
            Cminus1(j)= max(0,M_in-k-r21(j)+Cminus1(j-1));
        end
    end
```

```
    for j=1:n
        if Cplus(j)>Ucl;
            RL(i)=j-1;
            break
        elseif Cminus(j)<Lcl;
            RL(i)=j-1;
            break
        else
            RL(i)=n;
        end
    end
end

ARL=sum(RL)/Nsim
```