USING THE EXPONENTIALLY WEIGHTED MOVING AVERAGE $\overline{X}$ CHART FOR CHANGE POINT ANALYSIS

by

CLIFFORD N JONES

(Under the Direction of Charles W .Champ)

ABSTRACT

The exponentially weighted moving average (EWMA) chart $\overline{X}$ for monitoring a process mean was introduced by Roberts[1]. Various authors have studied this chart. Shamma and Shamma[17] introduced an EWMA of an EWMA chart, referred to as the double EWMA chart. Since then the EWMA of the EWMA of the EWMA (triple EWMA) chart and the EWMA of the EWMA of the EWMA of the EWMA (quadruple EWMA) charts have been introduced into the literature. Their claims are that the double EWMA $\overline{X}$ chart outperformed the EWMA chart for the monitoring the process mean. Further it was claimed that the triple EWMA $\overline{X}$ chart outperformed the double EWMA $\overline{X}$ chart and the quadruple EWMA chart outperformed the triple EWMA $\overline{X}$ chart. We demonstrate that an EWMA $\overline{X}$ chart can be designed that outperforms the double EWMA $\overline{X}$ chart. We show this using simulation. The EWMA $\overline{X}$ has been used to predict the change point. We provide a method using the likelihood function to predict the change point. A comparison is made with the Shewhart $\overline{X}$ chart.

INDEX WORDS: EWMA$^{(k)}$ Chart, Likelihood Function, Phase I, Phase II

USING THE EXPONENTIALLY WEIGHTED MOVING AVERAGE $\overline{X}$ CHART FOR CHANGE POINT ANALYSIS

by

CLIFFORD N JONES

B.S, Meru University Of Science And Technology, Kenya, 2016

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

©2023

CLIFFORD N JONES

All Rights Reserved

USING THE EXPONENTIALLY WEIGHTED MOVING AVERAGE $\overline{X}$ CHART FOR CHANGE POINT ANALYSIS

by

CLIFFORD N JONES

Major Professor: Charles W .Champ
Committee: Charles W .Champ
Divine Wanduku
Andrew Sills

Electronic Version Approved:
July, 2023

DEDICATION

This Thesis is dedicated to the entire family of Jones Kitavi.

ACKNOWLEDGMENTS

I would like to express my heartfelt appreciation to my esteemed advisor, Dr. Charles W. Champ, for his invaluable guidance and unwavering encouragement throughout my thesis journey. His insightful ideas and assistance have played a pivotal role in the successful completion of this work. I extend my sincere thanks to Dr. Andrew V. Sills and Dr. Divine Wanduku for their valuable suggestions and for graciously agreeing to be part of my committee. Their expertise and contributions have significantly elevated the quality of my research. I am deeply grateful to my family for their unwavering support and encouragement, which have served as a constant source of motivation in both my educational pursuits and personal endeavors. Lastly, I would like to commend the faculty, staff, and fellow graduates of the Department of Mathematics for their continuous support and camaraderie throughout my academic journey.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS 3

LIST OF TABLES 7

CHAPTER

1 INTRODUCTION 8

2 MODEL AND SAMPLING METHOD 9

2.1 Introduction 9

2.2 Process Model 9

2.3 Sampling Method 11

2.4 Conclusion 12

3 SOME NON-NORMAL DISTRIBUTIONS 13

3.1 Introduction 13

3.2 Family of Uniform Distributions 13

3.3 Family of Triangular Distributions 15

3.4 Family of Gamma Distributions 16

3.5 Family of Student's t-Distributions 17

3.6 A Family of Mixtures of Normal Distributions 18

3.7 Conclusion 19

4 EXPONENTIALLY WEIGHTED MOVING AVERAGE $\overline{X}$ CONTROL CHART 20

4.1 Introduction 20

4.2 Shewhart $\overline{X}$ Chart 20

4.3 EWMA $\overline{X}$ Chart for Monitoring the Process Mean 24

4.4 An Integral Equation Approach to the Run Length and Its Distribution 31

4.5 Robustness of the EWMA $\overline{X}$ Chart 39

4.6 Conclusion 40

5 EWMA$^{(k)}$ $\overline{X}$ CHARTS 41

5.1 Introduction 41

5.2 EWMA$^{(k)}$ $\overline{X}$ Chart with Known Parameters 42

5.3 EWMA$^{(2)}$ or Double EWMA $\overline{X}$ Chart 47

5.4 Triple EWMA $\overline{X}$ Chart 54

5.5 Quadruple EWMA $\overline{X}$ Chart 57

5.6 Design Methods 58

5.7 Conclusion 64

6 CHANGE POINT ANALYSIS 65

6.1 Introduction 65

6.2 Predicting the Change Point Using a Shewhart $\overline{X}$ Chart 65

6.3 Change Point Prediction Using a EWMA $\overline{X}$ Chart 73

6.4 Conclusion 83

7 CONCLUSIONS 84

7.1 General Conclusions 84

7.2 Areas for Further Research 84

REFERENCES 86

LIST OF TABLES

Table Page

5.1 Comparison of EWMA and DEWMA $\overline{X}$ Charts 60

5.2 Comparison of EWMA and DEWMA $\overline{X}$ Charts 61

5.3 Comparison of EWMA and DEWMA $\overline{X}$ Charts 62

5.4 Comparison of EWMA and DEWMA $\overline{X}$ Charts 63

6.1 In-control Known Paramters Case, ARL_0 equal to 200 69

6.2 In-control Estimated Parameters Case, ARL_0 equal to 200 73

6.3 In-control Known Parameters Case, ARL_0 equal to 200 78

6.4 In-control Estimated Parameters Case, ARL_0 equal to 200 82

CHAPTER 1

INTRODUCTION

Walter A. Shewhart introduced the quality control chart in the early 1920's. Shewart [18] introduced the concepts of a production process being in a state of in control and out of control in terms of types of variability found in the process. In Chapter 2, we discuss our model for a process being statistically in and out of control with respect to the Normal distribution. Chapter 3 discusses various non Normal distributions that one could use to study the robustness of a control chart to the assumption of the Normal distribution.

In Chapter 4, the exponentially weighted moving average (EWMA) $\overline{X}$ is discussed. Integral equations are derived that are useful in determining properties of the run length distribution. The double, triple, quadruple, etc. EWMA $\overline{X}$ are discussed in Chapter 5. These charts are referred to as the EWMA$^{(k)}$ $\overline{X}$ charts for $k = 2, 3, 4, \ldots$. A method is given of obtaining an EWMA $\overline{X}$ that outperforms an EWMA$^{(k)}$ $\overline{X}$ chart for $k = 2, 3, 4$.

Control charts have been used for predicting a change point in a production process. We discuss in Chapter 6 the use of the Shewhart $\overline{X}$ chart (an EWMA $\overline{X}$ chart with smoothing parameter $r = 1$) in predicting the change point using the likelihood function both with in-control parameters known and estimated. Further, we derived the likelihood function associated with the EWMA $\overline{X}$ chart and derive a method for predicting the change point with parameters known and estimated.

We give some concluding remarks along with some areas for futher research in our last chapter. References are given. This is followed by an Appendix containing some R programs used in this research.

CHAPTER 2

MODEL AND SAMPLING METHOD

2.1 Introduction

Duncan[3] point out three uses of control charts. (1) As an aid in removing initial assignable causes of variability. (2) As an aid in defining the meaning of the process being a state of statistical in control. (3) Monitoring for a change in the process. Both (1) and (2) are are performed in what is called a Phase I analysis of the process. We provide some estimation methods used in a Phase I analysis of the process. The monitoring phase is referred to as Phase II. Our interest is the use of control charts in the monitoring phase.

Most statistical methods are constructed assumimg a model for the data. Quality control charts are statistical methods. We discuss in this chapter our data model we will refer to as the independent Normal model. A commonly used simple model for the process being in-control and out-of-control is discussed. The sampling method we will use follows the method found in Shewhart[18] which he referred to as periodic sampling.

2.2 Process Model

We assume that the quality measurement X has a Normal distribution with mean μ and variance σ^2. We will use a simple model for the process being in a state of statistical in and out of control. The process is assumed to be in statistical in control $\mu = \mu_0$ and $\sigma^2 = \sigma_0^2$ or $\sigma = \sigma_0$, where μ_0 and σ_0^2 are possible values of μ and σ^2. If $\mu \neq \mu_0$ and/or $\sigma^2 \neq \sigma_0^2$, then the process is in a state of statistical out of control. Typically, the parameters μ_0 and σ_0^2 are unknown and will be estimated from a preliminary set of samples generated by the process when in a state of statistical in control. We assume the quality measurements taken on items from the process are independent. We refer to our model as the independent Normal model.

The standard Normal distribution was first described by Abraham de Moivre in 1733. The more general family of Normal distributions was first described by the German mathematician Carl Friedrich Gauss (1777-1885). Often a Normal distribution is referred to as a Gaussian distribution. The probability density function describing a member of the family of Normal distributions has the form

$$f\left(x\left|\mu,\sigma\right.\right)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^{2}},$$

where μ is the mean of the distribution and $\sigma > 0$ (σ^2) is the standard deviation (variance) of the distribution. The Normal distribution with $\mu = 0$ and $\sigma^2 = 1$ is called the standard Normal distribution. Its probability density $\phi\left(z\right)$ and cumulative distribution function $\Phi\left(z\right)$ describe the distribution of the transformation $Z = \left(\frac{X-\mu}{\sigma}\right)^2$. They have the forms

$$\phi\left(z\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^{2}} \text{ and } \Phi\left(z\right)=\int_{-\infty}^{z}\phi\left(t\right)dt.$$

Walter A. Shewhart developed the quality control chart as a statistical method. Shewhart[18] discusses the concepts of "natural" and "assignable" causes variability relative to an industrial process. A natural cause of variability is a cause of variability in a quality measurement(s) that have been designed into the process. Such a cause of variability can only be removed by redesigning the process. A process is said to be in control if it is operating with only natural causes of variability. Under our Normal model, we assume the process is a statistically in control state if $\mu = \mu_0$ and $\sigma = \sigma_0$. The in control values μ_0 and σ_0, respectively, are possible values of the process mean and process standard deviation that may or may not be known. An assignable cause of variability is a removable cause of variability in the process. When this is the case, we say the process is out of control. Under our model, we say the process is statistically out of control if $\mu \neq \mu_0$ and/or $\sigma \neq \sigma_0$ when an assignable cause is present in the process. In what is to follow, we assume the process is out-of-control only with respect to the process mean.

2.3 SAMPLING METHOD

In the Phase I of the process, we assume that the practitioner will have available m independent random samples each of size n from the output of an in control process. We represent the X measurements to be taken on these items by $X_{i,1}, \ldots, X_{i,n}$ for $i = 1, \ldots, m$. We assume the practitioner will use $\widehat{\mu}_0$ and $\widehat{\sigma}_0^2$, respectively, as estimators of μ_0 and σ_0^2 defined by

$$\widehat{\mu}_0 = \overline{\overline{X}}_0 = \frac{1}{m}\sum\nolimits_{i=1}^{m} \overline{X}_i \text{ and } \widehat{\sigma}_0^2 = \overline{V}_0 = \frac{1}{m}\sum\nolimits_{i=1}^{m} S_i^2,$$

where $\overline{X}_i$ and S_i^2 are respectively the mean and variance of the ith sample. Under our independent Normal model with $X \sim N(\mu_0, \sigma_0^2)$, the sampling distributions of $\widehat{\mu}_0 = \overline{\overline{X}}_0$ and $\widehat{\sigma}_0^2 = \overline{V}_0$ are determined as follows. Observe that $\overline{X}_i \sim N(\mu_0, \sigma_0^2/n)$ and $\widehat{\mu}_0 = \overline{\overline{X}}_0$ is a linear combination of the $\overline{X}_i$'s. It is well known that a linear combination of independent Normal random variables has a Normal distribution with mean

$$E\left(\overline{\overline{X}}_0\right) = E\left(\frac{1}{m}\sum\nolimits_{l=1}^{m} \overline{X}_i\right) = \frac{1}{m}\sum\nolimits_{l=1}^{m} E\left(\overline{X}_i\right) = \frac{1}{m}\sum\nolimits_{l=1}^{m} \mu_0 = \mu_0 \text{ and}$$
$$V\left(\overline{\overline{X}}_0\right) = V\left(\frac{1}{m}\sum\nolimits_{l=1}^{m} \overline{X}_i\right) = \frac{1}{m^2}\sum\nolimits_{l=1}^{m} V\left(\overline{X}_i\right) = \frac{1}{m^2}\sum\nolimits_{l=1}^{m} \frac{\sigma_0^2}{n} = \frac{\sigma_0^2}{mn}.$$

The distribution of $\widehat{\sigma}_0^2 = \overline{V}_0$ is obtained by observing that

$$\begin{aligned}\overline{V}_0 &= \frac{1}{m(n-1)}\sum\nolimits_{l=1}^{m} (n-1) S_i^2 = \frac{\sigma_0^2}{m(n-1)}\sum\nolimits_{l=1}^{m} \frac{(n-1) S_i^2}{\sigma_0^2} \\ &= \frac{\sigma_0^2}{m(n-1)}\sum\nolimits_{l=1}^{m} \chi_{i,n-1}^2,\end{aligned}$$

It is well known that

$$\chi^2_{i,n-1} = \frac{(n-1)\, S_i^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

and the sum of independent Chi square random variables is a Chi square distribution with degrees of freedom the sum of the degrees of freedom of the addends, then

$$\overline{V}_0 \sim \frac{\sigma_0^2}{m\,(n-1)} \sum\nolimits_{l=1}^{m} \chi^2_{i,n-1} = \frac{\sigma_0^2}{m\,(n-1)} \chi^2_{m(n-1)} \text{ or } \frac{m\,(n-1)\,\overline{V}_0}{\sigma_0^2} \sim \chi^2_{m(n-1)}.$$

Also note that the statistics $\overline{\overline{X}}_0$ and $\overline{V}_0$ are independent under our model.

In Phase II, the practitioner will periodically at time t (sample stage number) take the X measurements $X_{t,1}, \ldots, X_{t,n}$ on a sample of size n from the output of the process, for $t = 1, 2, 3, \ldots$. The mean of the sample to be taken at time t (sampling stage t) is denoted by $\overline{X}_t$. It is assumed the samples are independent random samples, for $t = 1, 2, 3, \ldots$. Under our independent Normal model, we have $\overline{X}_t \sim N(\mu, \sigma^2/n)$ are independent for $t = 1, 2, 3, \ldots$. If the process is statisticlly in control, then $\overline{X}_t \sim N(\mu_0, \sigma_0^2/n)$ for $t = 1, 2, 3, \ldots$.

2.4 Conclusion

Our process model and sampling methods has been discussed under which we will study the EWMA $\overline{X}$ and the EWMA$^{(k)}$ $\overline{X}$ charts. This includes our change point analysis.

CHAPTER 3

SOME NON-NORMAL DISTRIBUTIONS

3.1 INTRODUCTION

The EWMA $\overline{X}$ will be designed under the independent Normal model. In this chapter, we examine some non-Normal distributions to be used to study the robustness of the EWMA $\overline{X}$. These include the Uniform, Triangular, Gamma, and t-distributions. The Uniform and t-distributions and a sub family of the family of Triangular distributions are symetric about their mean. In the case of the Uniform and t-distributions, the tails of the distribution are higher than the Normal. The family of Gamma distributions and sub families of the Triangular distributions are skewed. The Gamma is skewed in the positive direction while a Triangular distribution can be selected to be skewed in either the positive or negative directions. Mixtured distributions are also discussed. A short discussion of each of these distributions is given in what follows.

3.2 FAMILY OF UNIFORM DISTRIBUTIONS

A random variable X has a Uniform distribution on the interval (a, b) with $a < b$ (or $X \sim UNIFORM\,(a, b)$) if the probability density and cumulative distribution functions are given by

$$f_X\left(x\,|a,b\right) = \frac{1}{b-a} I_{(a,b)}\left(x\right) \text{ and}$$
$$F_X\left(x\,|a,b\right) = \frac{x-a}{b-a} I_{(a,b)}\left(x\right) + I_{[b,\infty)}\left(x\right).$$

The indicator function $I_{\mathbf{A}}\left(x\right) = 1$ if $x \in \mathbf{A}$ and zero otherwise. A well known special case sets $a = 0$ and $b = 1$. The mean μ_X, variance σ_X^2, and standard deviation σ_X of the distribution of X are

$$\mu_X = \frac{a+b}{2}, \sigma_X^2 = \frac{(b-a)^2}{12}, \text{ and } \sigma_X = \frac{b-a}{\sqrt{12}}.$$

The $100\,(1-\alpha)$th percentile x_α is $x_\alpha = a + (b-a)\,(1-\alpha)$.

Theorem 3.1: The transformation

$$U = \frac{X-a}{b-a} \sim UNIFORM\,(0,1) \text{ or}$$

$$X = a + (b-a) \sim UNIFORM\,(a,b)\,.$$

Proof:

$$F_U\,(u) = P\,(U \le u)\,I_{(0,1)}\,(u) = P\left(\frac{X-a}{b-a} \le u\right) I_{(0,1)}\,(u)$$

$$= P\,(X \le a + (b-a)\,u)\,I_{(0,1)}\,(u)$$

$$= uI_{(0,1)}\,(u) + I_{[1,\infty)}\,(u)\,.$$

This is the cumulative distributon of a random variable with a Uniform distribution on the interval $(0,1)$.

Suppose that $U_1,\ldots,U_n$ are independent with a common Uniform distribution on the interval $(0,1)$, then the distribution of the sample mean $\overline{U}_n$ has probability density function

$$f_{\overline{U}_n}\,(\overline{u}) = \sum\nolimits_{k=0}^{n-1} \frac{n}{(n-1)!} \left(\sum\nolimits_{j=0}^{k} (-1)^j \binom{n}{j} (n\overline{u}-j)^{n-1}\right) I_{(k/n,(k+1)/n]}\,(\overline{u})\,.$$

These results can be found in Mood, Graybill, and Boes (1974). The mean and variance of $\overline{U}_n$ are

$$\mu_{\overline{U}_n} = \frac{a+b}{2} \text{ and } \sigma_{\overline{U}_n}^2 = \frac{(b-a)^2/12}{n}.$$

Suppose that $X_1,\ldots,X_n$ are independent with a common Uniform distribution on the interval (a,b), then

$$\overline{U}_n = a + (b-a)\,\overline{X}_n \text{ or } \overline{X}_n = \frac{\overline{U}_n - a}{b-a},$$

where

$$\overline{U}_n = \frac{1}{n} \sum\nolimits_{i=1}^{n} U_i \text{ and } \overline{X}_n = \frac{1}{n} \sum\nolimits_{i=1}^{n} X_i$$

with $U_i = (X_i - a) / (b - a)$. Thus, we have $\overline{X}_n = a + (b - a) \overline{U}_n$. It follows that the probability density function describing the distribution of $\overline{X}_n$ is

$$\begin{aligned} & f_{\overline{X}_n} (\overline{x} \,| a, b, n) \\ & = \frac{1}{b-a} \sum\nolimits_{k=0}^{n-1} \frac{n}{(n-1)!} \left[\sum\nolimits_{j=0}^{k} (-1)^j \binom{n}{j} (n (\overline{x} - a) / (b-a) - j)^{n-1} \right] \\ & \times I_{(a+(b-a)k/n, a+(b-a)(k+1)/n]} (\overline{x}), \end{aligned}$$

by observing that

$$\begin{aligned} k/n < \overline{u} \le (k+1)/n \text{ or} \\ a + (b-a) k/n < a + (b-a) \overline{u} \le a + (b-a)(k+1)/n \text{ or} \\ a + (b-a) k/n < \overline{x} \le a + (b-a)(k+1)/n. \end{aligned}$$

Several pseudo Uniform random number generators have been developed. Hull and Dobell[21] list references that discuss the statistical tests that have been used to test random number generators. In the R language, the pseudo Uniform random number generator is $runif\,(n, a, b)$. It generates n pseudo Uniform random numbers in the interval (a, b).

3.3 Family of Triangular Distributions

A simple nonnormal distribution is the triangular distribution. The general form of the probability density function is

$$f_X (x \,| a, b, c) = \left(\frac{2(x-a)}{(b-a)(c-a)} \right)^{I_{(a,b]}(x)} \left(\frac{2(c-x)}{(c-a)(c-b)} \right)^{I_{(b,c)}(x)} I_{(a,c)} (x)$$

for $a < b$ and $a \leq b \leq c$. It is easy to see that (1) $f_X(x\,|a,b,c) \geq 0$ for all real numbers x and (2)

$$\int_{-\infty}^{\infty} f_X(x\,|a,b,c)\,dx = \int_a^b \frac{2(x-a)}{(b-a)(c-a)}dx + \int_b^c \frac{2(c-x)}{(c-a)(c-b)}dx = 1.$$

For convenience, we write $X \sim TRI(a,b,c)$.

The cumulative distribution function $F_X(x\,|a,b,c)$ has the form

$$F_X(x\,|a,b,c) = \left(\frac{(x-a)^2}{(b-a)(c-a)}\right)^{I_{[a,b]}(x)} + \left(1 - \frac{(c-x)^2}{(c-a)(c-b)}\right)^{I_{[b,c]}(x)} + I_{[c,\infty)}(x).$$

It is shown in Bain and Engelhardt (1992) that $U = F_X(X\,|a,b,c)$ has Uniform distribution on the interval $(0,1)$. If U has a Uniform distribution on the interval $(0,1)$, then $F_X^{-1}(U\,|a,b,c)$ has the same distribution as X. The inverse of F_X is

$$\begin{aligned} U = F_X^{-1}(X\,|a,b,c) &= \left(a + \sqrt{(b-a)(c-a)U}\right)^{I_{(0,(b-a)/(c-a)]}(u)} \\ &+ \left(c - \sqrt{(c-a)(c-b)(1-U)}\right)^{I_{((b-a)/(c-a),1)}(u)}. \end{aligned}$$

Hence, the random variable $F_X^{-1}(U\,|a,b,c)$ has the same distribution as X if U is Uniform on the interval $(0,1)$.

3.4 Family of Gamma Distributions

A random variable X has a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ ($X \sim GAM(\alpha,\beta)$) if the distribution of X is described by the density function

$$f(x\,|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} I_{(0,\infty)}(x)$$

Consider the transformation $Y = 2X/\beta$. It is not difficult to show that $Y \sim \chi^2_{2\alpha}$. For a given α and β, one can use a random number generator that generates a Chi squared random variable $\chi^2_{2\alpha}$. The random variable $X = \beta\chi^2_{2\alpha}/2 \sim GAM\left(\alpha,\beta\right)$. It is not difficult to show that if $X_1,\ldots,X_n$ are independent with a common $GAM\left(\alpha,\beta\right)$ distribution, then the distribution of the sample mean $\overline{X}_n$ is a $GAM\left(\alpha\beta, \frac{\alpha\beta^2}{n}\right)$ distribution. Using the previous results, we have that the sample $\overline{X}_n = \alpha\beta^2\chi^2_{2\alpha\beta}/\left(2n\right)$ allowing one to generate a sample mean of n independent random variables each having a $GAM\left(\alpha,\beta\right)$ distribution. The R function $rgamma(n,\alpha,\beta)$ generates n random variables from a $GAM\left(\alpha,\beta\right)$ distribution.

3.5 Family of Student's t-Distributions

Suppose that the measurements $X_1,\ldots,X_n$ are independent with a common $N\left(\mu,\sigma^2\right)$ distribution. The mean $\overline{X}$ and variance S^2 of the sample are defined by

$$\overline{X} = \frac{1}{n}\sum\nolimits_{i=1}^{n} X_i \text{ and } S^2 = \frac{1}{n-1}\sum\nolimits_{i=1}^{n}\left(X_i-\overline{X}\right)^2.$$

The random variable

$$T = \frac{\overline{X}}{S/\sqrt{n}} = \frac{\left(\overline{X}-\mu\right)/\left(\sigma/\sqrt{n}\right)+\sqrt{n}\mu/\sigma}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/\left(n-1\right)}} = \frac{Z+\theta}{\sqrt{\chi^2_{n-1}/\left(n-1\right)}}$$

is said to have a noncentral t-distribution with $n-1$ degrees of freedom and noncentrality parameter $\theta = \sqrt{n}\mu/\sigma$, where $Z \sim N\left(0,1\right)$ and $W \sim \chi^2_{n-1}$ are independent. If $\theta = 0$, the distribution of T is said to have a central t-distribution with $n-1$ degrees of freedom. The mean and variance of the distribution of the noncentral T distribution are

$$\mu_T = \theta\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \text{and } \sigma^2_T = \frac{\left(n-1\right)\left(1+\theta^2\right)}{n-3} - \frac{\left(n-1\right)\theta^2}{2}\left(\frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right)^2.$$

The R language function $rt\left(n,df,ncp\right)$ can be used to generate n random variates from a non-central t-distribution with df degrees of freedom and non-centrality parameters found in the numeric vector ncp.

3.6 A Family of Mixtures of Normal Distributions

Let $f_{X_1}(x), \ldots, f_{X_n}(x)$ and $F_{X_1}(x), \ldots, F_{X_n}(x)$ be, respectfully, the probability density and cumulative distribution functions of the independent random variables $X_1, \ldots, X_n$. Define $f_Y(x)$ the probability density funtion describing the distribution of Y by

$$f_Y(y) = \sum_{i=1}^{n} p_i f_{X_i}(y),$$

where

$$p_i \geq 0 \text{ and} \sum_{i=1}^{n} p_i = 1.$$

The distribution of Y is called a finite mixture distribution. Note that $f_Y(x) \geq 0$ for all values of x and

$$\int_{-\infty}^{\infty} f_Y(y)\,dx = \int_{-\infty}^{\infty} \sum_{i=1}^{n} p_i f_{X_i}(y)\,dy = \sum_{i=1}^{n} p_i \int_{-\infty}^{\infty} f_{X_i}(y)\,dy$$
$$= \sum_{i=1}^{n} p_i \int_{-\infty}^{\infty} f_{X_i}(y)\,dy = \sum_{i=1}^{n} p_i = 1,$$

if at least one of the p_i's is greater than zero. Hence, the function $f_Y(y)$ meets the first two requirements to be a probability density function. If it describes the distribution of a random variable Y, then it is a density function. The cumulative distribution function $F_Y(y)$ is given by

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)\,dt = \sum_{i=1}^{n} p_i \int_{-\infty}^{y} f_{X_i}(t)\,dt = \sum_{i=1}^{n} p_i F_{X_i}(y).$$

It follows that

$$U = F_Y(Y) = \sum_{i=1}^{n} p_i F_{X_i}(Y) = \sum_{i=1}^{n} p_i U_i \sim UNIFORM\,(0,1)$$

with $U_i = F_{X_i}(Y) \sim UNIFORM\,(0,1)$.

Suppose that $X_1 \sim N(\mu_1, \sigma_1^2), \ldots, X_n \sim N(\mu_n, \sigma_n^2)$ are independent. We have that

$$f_Y(y) = \sum_{i=1}^{n} p_i \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\left(\frac{y-\mu_i}{\sigma_i}\right)^2}.$$

For the case in which $n = 2$, we have

$$\begin{aligned} f_Y(y) &= p\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{y-\mu_1}{\sigma_1}\right)^2} + (1-p)\frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2} \\ &= p f_{N\left(\mu_1,\sigma_1^2\right)}(y) + (1-p) f_{N\left(\mu_2,\sigma_2^2\right)}(y) \text{ and} \\ F_Y(y) &= p\int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{t-\mu_1}{\sigma_1}\right)^2} dt + (1-p)\int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}\left(\frac{t-\mu_2}{\sigma_2}\right)^2} dt \\ &= p F_{N\left(\mu_1,\sigma_1^2\right)}(y) + (1-p) F_{N\left(\mu_2,\sigma_2^2\right)}(y). \end{aligned}$$

It follows that

$$\begin{aligned} U = F_Y(Y) &= p F_{N\left(\mu_1,\sigma_1^2\right)}(Y) + (1-p) F_{N\left(\mu_2,\sigma_2^2\right)}(Y) \\ &= pU_1 + (1-p) U_2 \sim UNIFORM(0,1). \end{aligned}$$

3.7 CONCLUSION

Many statistical method are constructed under the assumption that a measurement follows a Normal distribution. We have discussed various non Normal distributions that can be used in the study of the robustness of a statistical method to the Normal model.

CHAPTER 4

EXPONENTIALLY WEIGHTED MOVING AVERAGE $\overline{X}$ CONTROL CHART

4.1 INTRODUCTION

Roberts[1]introduced the exponentially weighted moving average (EWMA) chart as the Geometric Moving Average (GMA) chart. The chart was designed to monitor for a change in the process mean μ of a quality measurement X. The EWMA chart has been discussed by a variety of authors(see Montgomery[15])). Champ,Woodal and Mohsen[8] define a family of cumulative sum type charts that include the family of EWMA charts.Jones, Champ, and Rigdon[10] studied the EWMA $\overline{X}$ chart with estimated parameters.

In the next section, the EWMA $\overline{X}$ chart is defined. A standardized version of the chart is defined that is useful in analyzing the run length distribution of the chart. In Section 3, integral equations are derived that are useful in analyzing the run length distribution of the chart.

4.2 SHEWHART $\overline{X}$ CHART

In the monitoring phase (Phase II), a Shewhart $\overline{X}$ chart is a plot of the points with coordinates $\left(t, \overline{X}_t\right)$ for $t = 1, 2, 3, \ldots$, where $\overline{X}_t$ is the mean of the tth sample. It is not difficult to show under our model that $E\left(\overline{X}_t\right) = \mu$ and $V\left(\overline{X}_t\right) = \sigma^2/n$. If the in-control parameters are known, then the lower and upper control limits for the chart are

$$LCL = \mu_0 - h\frac{\sigma_0}{\sqrt{n}} \text{ and } UCL = \mu_0 + h\frac{\sigma_0}{\sqrt{n}},$$

where $h > 0$. The chart signal a potential out-of-control process at sampling stage t if $\overline{X}_t \leq LCL$ or $\overline{X}_t \geq UCL$.

For the case in which the in-control process parameters μ_0 and σ_0 are estimated with, respectively, $\overline{\overline{X}}_0$ and $\overline{V}_0^{1/2}/c_{4,m,n}$, the control limits are

$$LCL = \overline{\overline{X}}_0 - h\frac{\overline{V}_0^{1/2}/c_{4,m}}{\sqrt{n}} \text{ and } UCL = \overline{\overline{X}}_0 + h\frac{\overline{V}_0^{1/2}/c_{4,m}}{\sqrt{n}}.$$

where

$$c_{4,m} = \frac{\sqrt{2}\Gamma\left(\frac{m(n-1)+1}{2}\right)}{\sqrt{m\left(n-1\right)}\Gamma\left(\frac{m(n-1)}{2}\right)}.$$

The control limits LCL and UCL are unbiased estimators of the control limits in the in-control parameters known case since $E\left(\overline{\overline{X}}_0\right) = \mu_0$ and $E\left(\overline{V}_0^{1/2}/c_{4,m}\right) = \sigma_0$.

The run length of the chart is defined as the sample number T at which the chart first signals. In the parameters known case, the run length T has a Geometric distribution with parameter p, where

$$p = 1 - P\left(LCL < \overline{X}_t < UCL\right).$$

We have assuming $\sigma = \sigma_0$

$$\begin{aligned}
p &= 1 - P\left(\mu_0 - h\frac{\sigma_0}{\sqrt{n}} < \overline{X}_t < \mu_0 + h\frac{\sigma_0}{\sqrt{n}}\right) = 1 - P\left(-h < \frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}} < h\right) \\
&= 1 - P\left(-h < \frac{\overline{X}_t - \mu + \mu - \mu_0}{\sigma_0/\sqrt{n}} < h\right) = 1 - P\left(-h < \frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}} < h\right) \\
&= 1 - P\left(-h < Z_t + \delta < h\right) = 1 - P\left(-h - \delta < Z_t < h - \delta\right) \\
&= 1 - \Phi\left(h - \delta\right) + \Phi\left(-h - \delta\right) = p\left(h, \delta\right),
\end{aligned}$$

where Φ is the cumulative distribution function of a standard Normal distribution,

$$Z_t = \frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} \sim N\left(0, 1\right) \text{ and } \delta = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}.$$

Hence, the average run length (ARL) and standard deviation of the run length (SRL) are

$$ARL = E\left(T\right) = \frac{1}{p\left(h, \delta\right)} \text{ and } SRL = \sqrt{V\left(T\right)} = \frac{\sqrt{1 - p\left(h, \delta\right)}}{p\left(h, \delta\right)}.$$

For a given value of k, the in-control ARL and SRL are $ARL = E(T) = 1/p(h,0)$ and $SRL = \sqrt{V(T)} = \sqrt{1-p(h,0)}/p(h,0)$. For example for $h = 3$, we have

$$p(3,0) = 1 - \Phi(3) + \Phi(-3) = 0.0026999344.$$

It follows that the in-control ARL and SRL are

$$\begin{aligned} ARL &= \frac{1}{0.0026999344} = 370.3793692 \text{ and} \\ SRL &= \frac{\sqrt{1-0.0026999344}}{0.0026999344} = 369.8790313 \end{aligned}$$

The run length distribution is a function of the random variables $\overline{\overline{X}}_0$ and $\overline{V}_0^{1/2}$ (see Chapter 2 for their definitions) as well as k, m, n, and δ. We have

$$P\left(T = t \,\middle|\, h, m, n, \delta, \overline{\overline{X}}_0, \overline{V}_0\right) = 1 - P\left(\overline{\overline{X}}_0 - h\frac{\overline{V}_0^{1/2}/c_{4,m}}{\sqrt{n}} < \overline{X}_t < \overline{\overline{X}}_0 + h\frac{\overline{V}_0^{1/2}/c_{4,m}}{\sqrt{n}}\right).$$

Note that assuming that $\sigma = \sigma_0$, we have

$$\begin{aligned} \frac{\overline{X}_t - \overline{\overline{X}}_0}{\left(\overline{V}_0^{1/2}/c_{4,m}\right)/\sqrt{n}} &= \frac{c_{4,m}}{\overline{V}_0^{1/2}/\sigma_0}\left(\frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}} - \frac{\overline{\overline{X}}_0 - \mu_0}{\sigma_0/\sqrt{mn}}/\sqrt{m}\right) \\ &= \frac{c_{4,m}}{\overline{W}_0^{1/2}/\sigma_0}\left(Z_t + \delta - \overline{Z}_0/\sqrt{m}\right), \end{aligned}$$

where

$$\begin{aligned} Z_t &= \frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}} \sim N(0,1), \delta = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}, \overline{Z}_0 = \frac{\overline{\overline{X}}_0 - \mu_0}{\sigma_0/\sqrt{mn}} \sim N(0,1), \text{ and} \\ \overline{W}_0 &= \frac{\overline{V}_0}{\sigma_0^2} \sim \chi^2_{m(n-1)}/\left[m(n-1)\right]. \end{aligned}$$

We see that

$$
\begin{aligned}
&P\left(T=t\left|h,m,n,\delta,\overline{\overline{X}}_0,\overline{V}_0\right.\right)\\
&1-P\left(-h<\frac{\overline{X}_t-\overline{\overline{X}}_0}{\left(\overline{V}_0^{1/2}/c_4\right)/\sqrt{n}}<h\left|\overline{\overline{X}}_0,\overline{V}_0\right.\right)\\
&=1-P\left(-h<\frac{c_{4,m}}{\overline{W}_0^{1/2}/\sigma_0}\left(Z_t+\delta-\overline{Z}_0/\sqrt{m}\right)<h\left|\overline{\overline{X}}_0,\overline{V}_0\right.\right)\\
&=1-P\left(\begin{array}{c}-c_{4,m}^{-1}\overline{W}_0^{1/2}h-\delta+\overline{Z}_0/\sqrt{m}<Z_t\\ <c_{4,m}^{-1}\overline{W}_0^{1/2}h-\delta+\overline{Z}_0/\sqrt{m}\left|\overline{\overline{X}}_0,\overline{V}_0\right.\end{array}\right)\\
&=1-\Phi\left(c_{4,m}^{-1}\overline{W}_0^{1/2}h-\delta+\overline{Z}_0/\sqrt{m}\right)-\Phi\left(-c_{4,m}^{-1}\overline{W}_0^{1/2}h-\delta+\overline{Z}_0/\sqrt{m}\right)\\
&=P\left(T=t\left|h,m,n,\delta,\overline{Z}_0,\overline{W}_0\right.\right).
\end{aligned}
$$

The conditional ARL and SRL given $\overline{Z}_0$ and $\overline{W}_0$ are

$$
\begin{aligned}
ARL\left(h,m,n,\delta\left|\overline{Z}_0,\overline{W}_0\right.\right)&=\frac{1}{P\left(T=t\left|h,m,n,\delta,\overline{Z}_0,\overline{W}_0\right.\right)}\text{ and}\\
SRL\left(h,m,n,\delta\left|\overline{Z}_0,\overline{W}_0\right.\right)&=\frac{\sqrt{1-P\left(T=t\left|h,m,n,\delta,\overline{Z}_0,\overline{W}_0\right.\right)}}{P\left(T=t\left|h,m,n,\delta,\overline{Z}_0,\overline{W}_0\right.\right)}.
\end{aligned}
$$

The unconditional probability $P\left(T=t\left|h,m,n,\delta,\overline{Z}_0,\overline{W}_0\right.\right)$ of a signal at any time t is

$$
P\left(T=t\left|h,m,n,\delta\right.\right)=\int_0^\infty\int_{-\infty}^\infty P\left(T=t\left|h,m,n,\delta,\overline{z}_0,\overline{w}_0\right.\right)f_{\overline{Z}_0}\left(\overline{z}_0\right)f_{\overline{W}_0}\left(\overline{w}_0\right)d\overline{z}_0d\overline{w}_0.
$$

The unconditional ARL and SRL are determined as follows.

$$
\begin{aligned}
ARL\left(h,m,n,\delta\right)&=\int_0^\infty\int_{-\infty}^\infty\frac{1}{p\left(h,m,n,\delta,\overline{z}_0,\overline{w}_0\right)}f_{\overline{Z}_0}\left(\overline{z}_0\right)f_{\overline{W}_0}\left(\overline{w}_0\right)d\overline{z}_0d\overline{w}_0\text{ and}\\
SRL\left(h,m,n,\delta\right)&=\int_0^\infty\int_{-\infty}^\infty\frac{\sqrt{1-p\left(h,m,n,\delta,\overline{z}_0,\overline{w}_0\right)}}{p\left(h,m,n,\delta,\overline{z}_0,\overline{w}_0\right)}f_{\overline{Z}_0}\left(\overline{z}_0\right)f_{\overline{W}_0}\left(\overline{w}_0\right)d\overline{z}_0d\overline{w}_0.
\end{aligned}
$$

Numerical integration can be used to determine the unconditional $P\left(T = t\,|h, m, n, \delta\right)$, $ARL\left(h, m, n, \delta\right)$, and $SRL\left(h, m, n, \delta\right)$. These values are not equal to their counterparts in the parameters known case.

4.3 EWMA $\overline{X}$ Chart for Monitoring the Process Mean

Under our model, the process is in-control if $\mu = \mu_0$ and $\sigma = \sigma_0$. First, we introduce the EWMA $\overline{X}$ chart under the assumption that μ_0 and σ_0 are known. In this case, the exponentially weighted moving average (EWMA) statistic U_t is defined as

$$U_0 = \mu_0 \text{ and } U_t = \left(1 - r\right) U_{t-1} + r\overline{X}_t,$$

where $0 < r \le 1$ for $t = 0, 1, 2, 3, \ldots$. The EWMA $\overline{X}$ chart is a plot of the points with coordinates (t, U_t) for $t = 1, 2, 3, \ldots$. The chart signals a potential out-of-control process at time t if $U_t \le LCL$ or $U_t \ge UCL$, where LCL and UCL are known respectively as the lower and upper control limits.

Theorem 4.3.1. If $0 < r < 1$, then for each positive integer t,

$$U_t = \left(1 - r\right)^t \mu_0 + r\sum\nolimits_{i=1}^{t} \left(1 - r\right)^{t-i} \overline{X}_i.$$

Proof: For $t = 1$, we have

$$\begin{aligned} U_1 &= \left(1 - r\right) U_0 + r\overline{X}_1 = \left(1 - r\right) \mu_0 + r\overline{X}_1 \\ &= \left(1 - r\right)^1 \mu_0 + r\sum\nolimits_{i=1}^{1} \left(1 - r\right)^{1-i} \overline{X}_i. \end{aligned}$$

Hence, the theorem is true for $t = 1$. Assume for $t > 1$ the theorem is true. We have

$$\begin{aligned} U_{t+1} &= (1-r)\,U_t + r\overline{X}_{t+1} \\ &= (1-r)\left[(1-r)^t\,\mu_0 + r\sum\nolimits_{i=1}^{t}(1-r)^{t-i}\,\overline{X}_i\right] + r\overline{X}_{t+1} \\ &= (1-r)^{t+1}\,\mu_0 + r\sum\nolimits_{i=1}^{t}(1-r)^{t+1-i}\,\overline{X}_i + r\overline{X}_{t+1} \\ &= (1-r)^{t+1}\,\mu_0 + r\sum\nolimits_{i=1}^{t+1}(1-r)^{t+1-i}\,\overline{X}_i. \end{aligned}$$

Hence, the theorem is true for $t+1$. By the Axiom of Induction, the theorem is true for all positive integers t.

Using the results of Theorem 4.3.1, the mean of the distribution of U_t is

$$\begin{aligned} E\left(U_t\right) &= (1-r)^t\,\mu_0 + r\sum\nolimits_{i=1}^{t}(1-r)^{t-i}\,E\left(\overline{X}_i\right) \\ &= \mu + (1-r)^t\left(\mu - \mu_0\right). \end{aligned}$$

If the process is in-control, then $E\left(U_t\right) = \mu_0$. The variance of the distribution of U_t is given by

$$\begin{aligned} V\left(U_t\right) &= r^2\sum\nolimits_{i=1}^{t}(1-r)^{2(t-i)}\,V\left(\overline{X}_i\right) = r^2\left[\sum\nolimits_{i=1}^{t}(1-r)^{2(t-i)}\right]\frac{\sigma^2}{n} \\ &= \frac{r^2\left[1-(1-r)^{2t}\right]}{1-(1-r)^2}\frac{\sigma^2}{n} = \frac{r\left[1-(1-r)^{2t}\right]}{2-r}\frac{\sigma^2}{n}. \end{aligned}$$

If $\sigma = \sigma_0$, then

$$V\left(U_t\right) = \frac{r^2\left[1-(1-r)^{2t}\right]}{1-(1-r)^2}\frac{\sigma_0^2}{n} = \frac{r\left[1-(1-r)^{2t}\right]}{2-r}\frac{\sigma_0^2}{n}.$$

Using these results, the variable control limits are defined by

$$\begin{aligned} LCL &= LCL_t = \mu_0 - h\sqrt{\frac{r\left[1-(1-r)^{2t}\right]}{2-r}}\frac{\sigma_0}{\sqrt{n}} \text{ and} \\ UCL &= UCL_t = \mu_0 + h\sqrt{\frac{r\left[1-(1-r)^{2t}\right]}{2-r}}\frac{\sigma_0}{\sqrt{n}}, \end{aligned}$$

where r and h are chart parameters Note that the control limits dependent on t.

Control limits that do not depend on t are known as fixed control limits. Note that

$$\lim_{t\to\infty} \frac{r^2\left[1-(1-r)^{2t}\right]}{1-(1-r)^2} = \frac{r^2}{1-(1-r)^2} = \frac{r}{2-r}.$$

We use these results to define fixed limits by

$$LCL = \mu_0 - h\sqrt{\frac{r}{2-r}}\frac{\sigma_0}{\sqrt{n}} \text{ and } UCL = \mu_0 + h\sqrt{\frac{r}{2-r}}\frac{\sigma_0}{\sqrt{n}}.$$

It follows that there are two types of EWMA charts. Those with variable limits that depend on t and those that have fixed limits that do not depend on t.

Theorem 4.3.2: If $r = 1$, then the EWMA $\overline{X}$ chart is a Shewhart $\overline{X}$ chart both in the in-control parameters known and estimated in-control parameters cases with chart parameter h.

Proof: We see that for $r = 1$, $U_t = (1-1)\,U_{t-1} + 1\overline{X}_t = \overline{X}_t$. Since the EWMA $\overline{X}$ chart is a plot of the points with coordinates $(t, U_t) = \left(t, \overline{X}_t\right)$, then we have a Shewhart $\overline{X}$ chart with chart parameter h.

Consider the transformation

$$U_t^* = \frac{U_t - \mu_0}{\sigma_0/\sqrt{n}}.$$

It follows that

$$\begin{aligned} U_0^* &= \frac{U_0 - \mu_0}{\sigma_0/\sqrt{n}} = \frac{\mu_0 - \mu_0}{\sigma_0/\sqrt{n}} = 0 \text{ and} \\ U_t^* &= \frac{(1-r)\,U_{t-1} + r\overline{X}_t - (1-r)\,\mu_0 - r\mu_0}{\sigma_0/\sqrt{n}} \\ &= \frac{(1-r)\,(U_{t-1} - \mu_0) + r\left(\overline{X}_t - \mu_0\right)}{\sigma_0/\sqrt{n}} \\ &= (1-r)\,\frac{U_{t-1} - \mu_0}{\sigma_0/\sqrt{n}} + r\frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}} \\ &= (1-r)\,U_{t-1}^* + r\frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}}, \end{aligned}$$

for $t > 0$. Assuming $\sigma = \sigma_0$, observe that

$$\frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}} = \frac{\overline{X}_t - \mu + \mu - \mu_0}{\sigma_0/\sqrt{n}} = \frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}} = Z_t + \delta,$$

where

$$Z_t = \frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} \sim N\,(0, 1) \text{ and } \delta = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}.$$

Note that the process is in a state of statistical in control with respect to the process mean μ if $\mu = \mu_0$ or $\delta = 0$. We then see that the stochastic behavior of the EWMA under our independent Normal model is described by the sequence of EWMA statistic U_t^* defined by

$$U_0^* = 0 \text{ and } U_t^* = (1-r)\,U_{t-1}^* + r\,(Z_t + \delta)\,.$$

The chart signals at time t if $U_t^* \leq LCL^*$ or $U_t^* \geq UCL^*$, where the variable control limits are

$$\begin{aligned} LCL^* &= LCL_t^* = -h\sqrt{\frac{r\left[1 - (1-r)^{2t}\right]}{2-r}} \text{ and} \\ UCL^* &= UCL_t^* = h\sqrt{\frac{r\left[1 - (1-r)^{2t}\right]}{2-r}} \end{aligned}$$

or with fixed control limits

$$LCL^* = -h\sqrt{\frac{r}{2-r}} \text{ and } UCL^* = h\sqrt{\frac{r}{2-r}}.$$

The EWMA chart that is a plot of the points with coordinates (t, U_t^*) cannot be observed. However, this chart is equivalent to EWMA chart that plots the points with coordinates (t, U_t) in the sense that either both charts signal or both charts do not signal at time t. The chart based on the EWMA statistics U_t^* is useful in determining run length properties of the chart.

An estimated parameters version of the EWMA $\overline{X}$ chart plots the points with coordinates (t, U_t) with

$$U_0 = \overline{\overline{X}}_0 \text{ and } U_t = (1-r)\, U_{t-1} + r\overline{X}_t.$$

The variable control limits for this chart are

$$LCL_t = \overline{\overline{X}}_0 - h\sqrt{\frac{r\left[1-(1-r)^{2t}\right]}{2-r}}\frac{\overline{V}_0^{1/2}/c_4}{\sqrt{n}} \text{ and}$$
$$UCL_t = \overline{\overline{X}}_0 + h\sqrt{\frac{r\left[1-(1-r)^{2t}\right]}{2-r}}\frac{\overline{V}_0^{1/2}/c_4}{\sqrt{n}}.$$

The fixed control limits are

$$LCL = \overline{\overline{X}}_0 - h\sqrt{\frac{r}{2-r}}\frac{\overline{V}_0^{1/2}/c_4}{\sqrt{n}} \text{ and } UCL = \overline{\overline{X}}_0 + h\sqrt{\frac{r}{2-r}}\frac{\overline{V}_0^{1/2}/c_4}{\sqrt{n}}.$$

These limits are unbiased estimators of the control limits for the in-control parameters known case.

Consider the transformation

$$U_t^* = \frac{U_t - \overline{\overline{X}}_0}{\left(\overline{V}_0/c_{4,m}\right)/\sqrt{n}}.$$

We see that

$$U_t^* = (1-r)\,\frac{U_{t-1} - \overline{\overline{X}}_0}{\left(\overline{V}_0/c_4\right)/\sqrt{n}} + r\frac{\overline{X}_t - \overline{\overline{X}}_0}{\left(\overline{V}_0/c_4\right)/\sqrt{n}}.$$

Defining

$$Y_t^* = \frac{c_{m,n}}{c_{4,m}}\frac{\overline{X}_t - \overline{\overline{X}}_0}{\left(\overline{V}_0/c_{4,m}\right)/\sqrt{n}} = c_{m,n}\frac{\overline{X}_t - \overline{\overline{X}}_0}{\overline{V}_0/\sqrt{n}},$$

where

$$c_{m,n} = \frac{\sqrt{2}\Gamma\left(\frac{m(n-1)}{2}\right)}{\sqrt{m\,(n-1)}\Gamma\left(\frac{m(n-1)-1}{2}\right)}.$$

Note that the conditional expectation of Y_t^* given $\overline{X}_t$ is

$$E\left(Y_t^*\,|\overline{X}_t\right) = E\left(c_{m,n}\frac{\overline{X}_t - \overline{\overline{X}}_0}{\overline{V}_0/\sqrt{n}}\,|\overline{X}_t\right) = \frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}}.$$

To see this, observe that

$$\begin{aligned} E\left(\frac{\overline{X}_t - \overline{\overline{X}}_0}{\overline{V}_0^{1/2}/\sqrt{n}}\,|\overline{X}_t\right) &= \sqrt{n}E\left(\overline{X}_t - \overline{\overline{X}}_0\right)E\left(\overline{V}_0^{-1/2}\right) \\ &= \sqrt{n}\left(\overline{X}_t - \mu_0\right)E\left(\overline{V}_0^{-1/2}\right). \end{aligned}$$

Recall that

$$\overline{V}_0 \sim \frac{\sigma_0^2}{m\,(n-1)}\chi^2_{m(n-1)} \text{ or } \overline{V}_0^{-1/2} \sim \frac{\sqrt{m\,(n-1)}}{\sigma_0}\left(\chi^2_{m(n-1)}\right)^{-1}.$$

We have

$$E\left(\overline{V}_0^{-1/2}\right) = \frac{\sqrt{m(n-1)}}{\sigma_0} E\left[\left(\chi^2_{m(n-1)}\right)^{-1}\right]$$
$$= \frac{\sqrt{m(n-1)}}{\sigma_0}\int_0^\infty w^{-1/2}\frac{1}{\Gamma\left(\frac{m(n-1)}{2}\right)2^{m(n-1)/2}} w^{m(n-1)/2-1}e^{-w/2}dw$$
$$= \frac{\sqrt{m(n-1)}\Gamma\left(\frac{m(n-1)-1}{2}\right)2^{[m(n-1)-1]/2}}{\Gamma\left(\frac{m(n-1)}{2}\right)2^{m(n-1)/2}}\sigma_0^{-1}$$
$$\times\int_0^\infty \frac{1}{\Gamma\left(\frac{m(n-1)-1}{2}\right)2^{[m(n-1)-1]/2}} w^{[m(n-1)-1]/2-1}e^{-w/2}dw$$
$$= \frac{\sqrt{m(n-1)}\Gamma\left(\frac{m(n-1)-1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{m(n-1)}{2}\right)}\sigma_0^{-1} = c_{m,n}^{-1}\sigma_0^{-1}.$$

Hence,

$$E\left(\frac{\overline{X}_t - \overline{\overline{X}}_0}{\overline{V}_0^{1/2}/\sqrt{n}}\left|\overline{X}_t\right.\right) = \sqrt{n}\left(\overline{X}_t - \mu_0\right)c_{m,n}^{-1}\sigma_0^{-1} = c_{m,n}^{-1}\frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}} \text{ or}$$
$$E\left(Y_t^*\left|\overline{X}_t\right.\right) = \frac{\overline{X}_t - \mu_0}{\sigma_0/\sqrt{n}}.$$

An estimated parameters version of the chart plots the points with coordinates (t, U_t^*) with $U_0^* = 0$ and

$$U_t^* = (1-r)U_{t-1}^* + rY_t^*.$$

The chart signals at sampling stage t if $U_t^* \leq LCL^*$ or $U_t^* \geq UCL^*$, where the variable control limits for this chart are

$$LCL_t^* = -h\sqrt{\frac{r\left[1-(1-r)^{2t}\right]}{2-r}} \text{ and } UCL_t^* = h\sqrt{\frac{r\left[1-(1-r)^{2t}\right]}{2-r}}.$$

The fixed control limits are

$$LCL^* = -h\sqrt{\frac{r}{2-r}} \text{ and } UCL^* = \overline{\overline{X}}_0 + h\sqrt{\frac{r}{2-r}}.$$

Assuming $\sigma = \sigma_0$, note that Y_t^* can be expressed as

$$Y_t^* = \frac{c_{m,n}}{\overline{V}_0^{1/2}/\sigma_0}\left[\frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}} - \frac{\overline{\overline{X}}_0 - \mu_0}{\sigma_0/\sqrt{mn}}/\sqrt{m}\right]$$
$$= c_{m,n}\overline{W}_0^{-1/2}\left(Z_t + \delta - \overline{Z}_0/\sqrt{m}\right),$$

where

$$\overline{W}_0 = \frac{\overline{V}_0}{\sigma_0^2} \sim \chi^2_{m(n-1)}/\left[m\left(n-1\right)\right], Z_t = \frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} \sim N\left(0,1\right),$$
$$\delta = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}, \text{ and } \overline{Z}_0 = \frac{\overline{\overline{X}}_0 - \mu_0}{\sigma_0/\sqrt{mn}} \sim N\left(0,1\right).$$

In summary, we have $U_0^* = 0$ and

$$U_t^* = \left(1-r\right)U_{t-1}^* + rc_{m,n}\overline{W}_0^{-1/2}\left(Z_t + \delta - \overline{Z}_0/\sqrt{m}\right).$$

4.4 An Integral Equation Approach to the Run Length and Its Distribution

The run length T is the number of the Phase II sample in which the chart first signals. The run length is a discrete random variable with support the positive integers. Suppose in the in-control parameters known case, we give the EWMA $\overline{X}$ chart based on U_t^* a head-start such that

$$U_0^* = u \text{ and } U_t^* = \left(1-r\right)U_{t-1}^* + r\left(Z_t + \delta\right),$$

with variable control limits

$$LCL_t = -h\sqrt{\frac{r\left[1-\left(1-r\right)^{2t}\right]}{2-r}} \text{ and}$$
$$UCL_t = h\sqrt{\frac{r\left[1-\left(1-r\right)^{2t}\right]}{2-r}},$$

where $-h\sqrt{r} < u < h\sqrt{r}$. Define the function $pr^*(t\,|u)$ by

$$pr^*(t\,|u) = P(T = t\,|U_0^* = u).$$

This is the probability mass function that describes the run length distribution of the chart based on U_t^* with a head-start u (see Lucas and Saccucci[12])

Note that for $t = 1$,

$$LCL_1 = -h\sqrt{r} \text{ and } UCL_1 = h\sqrt{r}.$$

Observe that

$$\begin{aligned}
& pr^*(1\,|u) \\
&= P(T = 1\,|U_0^* = u) \\
&= 1 - P\left(-h\sqrt{r} < U_1^* < h\sqrt{r}\,|U_0^* = u\right) \\
&= 1 - P\left(-h\sqrt{r} < (1-r)\,u + r\,(Z_1 + \delta) < h\sqrt{r}\right) \\
&= 1 - P\left(\frac{-h\sqrt{r} - (1-r)\,u - r\delta}{r} < Z_1 < \frac{h\sqrt{r} - (1-r)\,u - r\delta}{r}\right) \\
&= 1 - \Phi\left(\frac{h\sqrt{r} - (1-r)\,u - r\delta}{r}\right) + \Phi\left(\frac{-h\sqrt{r} - (1-r)\,u - r\delta}{r}\right).
\end{aligned}$$

For $t > 1$, we have

$$\begin{aligned}
pr^*(t\,|u) &= P(T = 1\,|U_0^* = u) \\
&= P\left(T = t, -h\sqrt{r} < U_1^* < h\sqrt{r}\,|U_0^* = u\right) \\
&= P\left(T = t\,\middle|U_0^* = u, -h\sqrt{r} < U_1^* < h\sqrt{r}\right) \\
&\times P\left(-h\sqrt{r} < U_1^* < h\sqrt{r}\,|U_0^* = u\right) \\
&= \int_{-h\sqrt{r}}^{h\sqrt{r}} P(T - 1 = t - 1\,|U_0^* = u, U_1^* = u_1^*) \\
&\times f_{U_1^*|U_0^*}(u_1^*\,|u)\,du_1^* \\
&= \int_{-h\sqrt{r}}^{h\sqrt{r}} pr^*(t - 1\,|u_1^*)\,f_{U_1^*|U_0^*}(u_1^*\,|u)\,du_1^*.
\end{aligned}$$

Note that the distribution of $T - 1$ is the same as the distribution of T if $U_0^* = u_1^*$.

Observe that

$$\begin{aligned}
F_{U_1^*|U_0^*=u}(u_1^*\,|u) &= P(U_1^* \leq u_1^*\,|U_0^* = u) = P((1-r)\,u + r(Z_1 + \delta) \leq u_1^*) \\
&= P\left(Z_1 \leq \frac{u_1^* - (1-r)\,u - r\delta}{r}\right) = \Phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right),
\end{aligned}$$

where $\Phi(z)$ is the cumulative distribution function of the standard Normal distribution. It follows that

$$f_{U_1^*|U_0^*=u}(u_1^*\,|u) = \frac{1}{r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{\lambda r}\right),$$

where $\phi(z)$ is the probability density function of the standard Normal distribution. Hence,

$$pr^*(t\,|u) = \int_{-h\sqrt{r}}^{h\sqrt{r}} pr^*(t - 1\,|u_1^*)\,\frac{1}{r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right)du_1^*.$$

In summary, we have

$$pr^*(t\,|u) = \begin{cases} 1 - \Phi\left(\frac{h\sqrt{r}-(1-r)u-r\delta}{\lambda r}\right) + \Phi\left(\frac{-h\sqrt{r}-(1-r)u-r\delta}{\lambda r}\right), & \text{for } t = 1; \\ \displaystyle\int_{-h\sqrt{r}}^{h\sqrt{r}} pr^*(t - 1\,|u_1^*)\,\frac{1}{\lambda r}\phi\left(\frac{u_1^*-(1-r)u-r\delta}{\lambda r}\right)du_1^*, & \text{for } t > 1. \end{cases}$$

Note that $pr^*(t\,|u)$ is also a function of δ and λ. We now has an interative method for obtaining the run length distribution. A numerical method would be used to evaluate for each t $pr^*(t\,|u)$. Errors can accumulate using this method.Woodal[19] gives a method for approximating the tail distribution of a cumulative sum (CUSUM) $\overline{X}$ chart. This method can be applied to the EWMA $\overline{X}$ chart. Woodall's method is also useful for obtaining percentiles of the run length distribution.

The average run length (ARL) is a function of u, δ, and λ. It can be expressed as

$$\begin{aligned}
ARL(\delta,\lambda,u) &= \sum\nolimits_{t=1}^{\infty} tpr^*(t\,|\delta,\lambda,u) \\
&= pr^*(1\,|\delta,\lambda,u) + \sum\nolimits_{t=2}^{\infty} tpr^*(t\,|\delta,\lambda,u) \\
&= 1-\Phi\left(\frac{h\sqrt{r}-(1-r)\,u-r\delta}{\lambda r}\right)+\Phi\left(\frac{-h\sqrt{r}-(1-r)\,u-r\delta}{\lambda r}\right) \\
&+ \sum\nolimits_{t=2}^{\infty} t\int_{-h\sqrt{r}}^{h\sqrt{r}} pr^*(t-1\,|u_1^*)\,\frac{1}{\lambda r}\phi\left(\frac{u_1^*-(1-r)\,u-r\delta}{\lambda r}\right)du_1^* \\
&= 1-\Phi\left(\frac{h\sqrt{r}-(1-r)\,u-r\delta}{\lambda r}\right)+\Phi\left(\frac{-h\sqrt{r}-(1-r)\,u-r\delta}{\lambda r}\right) \\
&+ \int_{-h\sqrt{r}}^{h\sqrt{r}}\left[\sum\nolimits_{t=1}^{\infty}(1+t)\,pr^*(t\,|u_1^*)\right]\frac{1}{\lambda r}\phi\left(\frac{u_1^*-(1-r)\,u-r\delta}{\lambda r}\right)du_1^* \\
&= 1-\Phi\left(\frac{h\sqrt{r}-(1-r)\,u-r\delta}{\lambda r}\right)+\Phi\left(\frac{-h\sqrt{r}-(1-r)\,u-r\delta}{\lambda r}\right) \\
&+ \int_{-h\sqrt{r}}^{h\sqrt{r}}\left[\sum\nolimits_{t=1}^{\infty}(1+t)\,pr^*(t\,|u_1^*)\right]\frac{1}{\lambda r}\phi\left(\frac{u_1^*-(1-r)\,u-r\delta}{\lambda r}\right)du_1^*
\end{aligned}$$

Noting that

$$\begin{aligned}
\sum\nolimits_{t=1}^{\infty}(1+t)\,pr^*(t\,|u_1^*) &= \sum\nolimits_{t=1}^{\infty} pr^*(t\,|u_1^*) + \sum\nolimits_{t=1}^{\infty} tpr^*(t\,|u_1^*) \\
&= 1 + ARL(\delta,\lambda,u_1^*),
\end{aligned}$$

we have

$$
\begin{aligned}
ARL\left(\delta,u\right) &= 1-\Phi\left(\frac{h\sqrt{r}-\left(1-r\right)u-r\delta}{r}\right)+\Phi\left(\frac{-h\sqrt{r}-\left(1-r\right)u-r\delta}{r}\right)\\
&+\int_{-h\sqrt{r}}^{h\sqrt{r}}\left[1+ARL\left(\delta,u_1^*\right)\right]\frac{1}{r}\phi\left(\frac{u_1^*-\left(1-r\right)u-r\delta}{r}\right)du_1^*\\
&= 1-\Phi\left(\frac{h\sqrt{r}-\left(1-r\right)u-r\delta}{r}\right)+\Phi\left(\frac{-h\sqrt{r}-\left(1-r\right)u-r\delta}{\lambda r}\right)\\
&+\Phi\left(\frac{h\sqrt{r}-\left(1-r\right)u-r\delta}{r}\right)-\Phi\left(\frac{-h\sqrt{r}-\left(1-r\right)u-r\delta}{r}\right)\\
&+\int_{-h\sqrt{r}}^{h\sqrt{r}}ARL\left(\delta,u_1^*\right)\frac{1}{r}\phi\left(\frac{u_1^*-\left(1-r\right)u-r\delta}{r}\right)du_1^*\\
&= 1+\int_{-h\sqrt{r}}^{h\sqrt{r}}ARL\left(\delta,u_1^*\right)\frac{1}{r}\phi\left(\frac{u_1^*-\left(1-r\right)u-r\delta}{r}\right)du_1^*.
\end{aligned}
$$

Consider the transformation $u_1^* = h\sqrt{r}y$. We have

$$
\begin{aligned}
&\int_{-h\sqrt{r}}^{h\sqrt{r}}ARL\left(\delta,u_1^*\right)\frac{1}{r}\phi\left(\frac{u_1^*-\left(1-r\right)u-r\delta}{r}\right)du_1^*\\
&=\int_{-1}^{1}ARL\left(\delta,h\sqrt{r}y\right)\frac{h\sqrt{r}}{r}\phi\left(\frac{h\sqrt{r}y-\left(1-r\right)u-r\delta}{r}\right)dy.
\end{aligned}
$$

A Gaussian numerical integration method using Legendre polynomials can now be used to obtain the numeric value of the integral

$$
\begin{aligned}
&\int_{-1}^{1}ARL\left(\delta,h_1y\right)\frac{h\sqrt{r}}{r}\phi\left(\frac{h\sqrt{r}y-\left(1-r\right)u-r\delta}{r}\right)dy\\
&\approx\sum\nolimits_{j=1}^{\eta}ARL\left(\delta,h_1y_j\right)\frac{h\sqrt{r}}{r}\phi\left(\frac{h\sqrt{r}y_j-\left(1-r\right)u-r\delta}{r}\right)w_i
\end{aligned}
$$

with nodes and weights (y_i, w_i) for $i = 1, \ldots, \eta$. This gives us a system of approximate equations

$$
ARL\left(\delta,y_i\right)\approx 1+\sum\nolimits_{j=1}^{\eta}ARL\left(\delta,\lambda,h_1y_j\right)\frac{h\sqrt{r}}{\lambda r}\phi\left(\frac{h\sqrt{r}y_j-\left(1-r\right)y_i-r\delta}{\lambda r}\right)w_i
$$

for $i = 1, \ldots, \eta$. This system can be expressed using matrix notation by

$$
\mathbf{m} = \mathbf{1}+\mathbf{Q}\mathbf{m} \text{ or } \mathbf{m} = \left(\mathbf{I}-\mathbf{Q}\right)^{-1}\mathbf{1},
$$

where I is the $\eta \times \eta$ identity matrix,

$$\mathbf{m} = \begin{bmatrix} ARL\left(\delta, y_1\right) \\ ARL\left(\delta, y_2\right) \\ \vdots \\ ARL\left(\delta, y_\eta\right) \end{bmatrix}, \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} Q_{11} & Q_{12} & \dots & Q_{1\eta} \\ Q_{21} & Q_{22} & \dots & Q_{2\eta} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{\eta 1} & Q_{\eta 2} & \dots & Q_{\eta\eta} \end{bmatrix}, \text{ and}$$

$$Q_{ij} = \frac{1}{r}\phi\left(\frac{u_j - (1-r)\,u_i - r\delta}{r}\right) w_j.$$

The variance of $VRL\left(\delta, u\right)$ of the run length distribution can be determined by

$$\begin{aligned} VRL\left(\delta, u\right) &= \sum\nolimits_{t=1}^{\infty} \left(t - ARL\left(\delta, u\right)\right)^2 pr^*\left(t\left|\delta, u\right.\right) \\ &= \sum\nolimits_{t=1}^{\infty} t^2 pr^*\left(t\left|\delta, u\right.\right) - \left[ARL\left(\delta, u\right)\right]^2 \\ &= E\left(T^2\left|\delta, u\right.\right) - \left[ARL\left(\delta, u\right)\right]^2. \end{aligned}$$

Observe that

$$
\begin{aligned}
E\left(T^2\left|\delta, u\right.\right) &= \sum\nolimits_{t=1}^{\infty} t^2 pr^*\left(t\left|\delta, \lambda, u\right.\right) \\
&= pr^*\left(1\left|\delta, u\right.\right) + \sum\nolimits_{t=2}^{\infty} t^2 pr^*\left(t\left|\delta, u\right.\right) \\
&= pr^*\left(1\left|\delta, u\right.\right) + \sum\nolimits_{t=1}^{\infty} \left(1+t\right)^2 pr^*\left(1+t\left|\delta, u\right.\right) \\
&= pr^*\left(1\left|\delta, u\right.\right) + \sum\nolimits_{t=1}^{\infty} \left(1+2t+t^2\right) pr^*\left(1+t\left|\delta, u\right.\right) \\
&= pr^*\left(1\left|\delta, u\right.\right) + \sum\nolimits_{t=2}^{\infty} pr^*\left(t\left|\delta, u\right.\right) \\
&+ 2\sum\nolimits_{t=1}^{\infty} tpr^*\left(1+t\left|\delta, u\right.\right) + \sum\nolimits_{t=1}^{\infty} t^2 pr^*\left(1+t\left|\delta, u\right.\right) \\
&= 1 + 2\sum\nolimits_{t=1}^{\infty} t\int_{-h\sqrt{r}}^{h\sqrt{r}} pr^*\left(t\left|\delta, u_1^*\right.\right)\frac{1}{\lambda r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^* \\
&+ \sum\nolimits_{t=1}^{\infty} t^2\int_{-h\sqrt{r}}^{h\sqrt{r}} pr^*\left(t\left|\delta, u_1^*\right.\right)\frac{1}{r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^* \\
&= 1 + 2\int_{-h\sqrt{r}}^{h\sqrt{r}} \sum\nolimits_{t=1}^{\infty} tpr^*\left(t\left|\delta, u_1^*\right.\right)\frac{1}{\lambda r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^* \\
&+ \int_{-h\sqrt{r}}^{h\sqrt{r}} \sum\nolimits_{t=1}^{\infty} t^2 pr^*\left(t\left|\delta, u_1^*\right.\right)\frac{1}{\lambda r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^* \\
&= 1 + 2\int_{-h\sqrt{r}}^{h\sqrt{r}} ARL\left(\delta, u_1^*\right)\frac{1}{\lambda r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^* \\
&+ \int_{-h\sqrt{r}}^{h\sqrt{r}} E\left(T^2\left|\delta, u_1^*\right.\right)\frac{1}{\lambda r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^*.
\end{aligned}
$$

In summary, we have

$$
\begin{aligned}
E\left(T^2\left|\delta, u\right.\right) = 1 &+ 2\int_{-h\sqrt{r}}^{h\sqrt{r}} ARL\left(\delta, u_1^*\right)\frac{1}{\lambda r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^* \\
&+ \int_{-h\sqrt{r}}^{h\sqrt{r}} E\left(T^2\left|\delta, u_1^*\right.\right)\frac{1}{\lambda r}\phi\left(\frac{u_1^* - (1-r)\,u - r\delta}{r}\right) du_1^*.
\end{aligned}
$$

In the parameters estimated case, we define

$$
pr^*\left(t\left|u, \overline{z}_0, \overline{v}_0\right.\right) = P\left(T = t\left|U_0^* = u, \overline{Z}_0 = \overline{z}_0, \overline{V}_0 = \overline{v}_0\right.\right).
$$

It is not difficult to show that

$$pr^*\left(1\left|u,\overline{z}_0,\overline{v}_0\right.\right)=1-\Phi\left(\overline{w}_0^{1/2}\frac{h\sqrt{r}-(1-r)\,u-r\delta+\overline{z}_0/\sqrt{m}}{rc_{m,n}}\right)$$
$$+\,\Phi\left(\frac{-h\sqrt{r}-(1-r)\,u-r\delta+\overline{z}_0/\sqrt{m}}{rc_{m,n}\overline{w}_0^{-1/2}}\right).$$

For $t>1$, one can show that

$$pr^*\left(t\left|u,\overline{z}_0,\overline{v}_0\right.\right)=\int_{-h\sqrt{r}}^{h\sqrt{r}}pr^*\left(t-1\left|u_1^*,\overline{z}_0,\overline{v}_0\right.\right)$$
$$\times\, f_{U_1^*|U_0^*}\left(u_1^*\left|u,\overline{z}_0,\overline{v}_0\right.\right)du_1^*,$$

with

$$f_{U_1^*|U_0^*}\left(u_1^*\left|u\right.\right)=\frac{1}{r}\phi\left(\overline{w}_0^{1/2}\frac{u_1^*-(1-r)\,u-r\delta+\overline{z}_0/\sqrt{m}}{rc_{m,n}}\right).$$

In summary, we have

$$pr^*\left(t\left|u,\overline{z}_0,\overline{v}_0\right.\right)=\begin{cases}1-\Phi\left(\overline{w}_0^{1/2}\frac{h\sqrt{r}-(1-r)u-r\delta+\overline{z}_0/\sqrt{m}}{rc_{m,n}}\right)\\+\Phi\left(\frac{-h\sqrt{r}-(1-r)u-r\delta+\overline{z}_0/\sqrt{m}}{rc_{m,n}\overline{w}_0^{-1/2}}\right),\\ \text{for } t=1;\\ \int_{-h\sqrt{r}}^{h\sqrt{r}}pr^*\left(t-1\left|u_1^*,\overline{z}_0,\overline{v}_0\right.\right)\\ \times\frac{1}{r}\phi\left(\overline{w}_0^{1/2}\frac{u_1^*-(1-r)u-r\delta+\overline{z}_0/\sqrt{m}}{rc_{m,n}}\right)du_1^*,\\ \text{for } t>1.\end{cases}$$

The conditional ARL is the exact solutions to the integral equation

$$ARL\left(\delta,u,\overline{z}_0,\overline{v}_0\right)=1+\int_{-h\sqrt{r}}^{h\sqrt{r}}ARL\left(\delta,u_1^*,\overline{z}_0,\overline{v}_0\right)$$
$$\times\,\frac{1}{r}\phi\left(\overline{w}_0^{1/2}\frac{u_1^*-(1-r)\,u-r\delta+\overline{z}_0/\sqrt{m}}{rc_{m,n}}\right)du_1^*.$$

The conditional variance of the run length distribution

$$VRL\left(\delta,u,\overline{z}_0,\overline{v}_0\right)=E\left(T^2\left|\delta,u,\overline{z}_0,\overline{v}_0\right.\right)-\left[ARL\left(\delta,u,\overline{z}_0,\overline{v}_0\right)\right]^2.$$

The unconditional probability mass function, average run length, and variance of the run length are determined as follows. We have

$$pr^{*}\left(t\left|u\right.\right)=\int_{0}^{\infty}\int_{-\infty}^{\infty}pr^{*}\left(t\left|u,\overline{z}_{0},\overline{v}_{0}\right.\right)f_{\overline{Z}_{0},\overline{W}_{0}}\left(\overline{z}_{0},\overline{w}_{0}\right)d\overline{z}_{0}d\overline{w}_{0};$$

$$ARL\left(\delta,u\right)=\int_{0}^{\infty}\int_{-\infty}^{\infty}ARL\left(\delta,u,\overline{z}_{0},\overline{v}_{0}\right)f_{\overline{Z}_{0},\overline{W}_{0}}\left(\overline{z}_{0},\overline{w}_{0}\right)d\overline{z}_{0}d\overline{w}_{0};\text{ and}$$

$$VRL\left(\delta,u\right)=\int_{0}^{\infty}\int_{-\infty}^{\infty}VRL\left(\delta,u,\overline{z}_{0},\overline{v}_{0}\right)f_{\overline{Z}_{0},\overline{W}_{0}}\left(\overline{z}_{0},\overline{w}_{0}\right)d\overline{z}_{0}d\overline{w}_{0},$$

where $f_{\overline{Z}_{0},\overline{W}_{0}}\left(\overline{z}_{0},\overline{w}_{0}\right)=f_{\overline{Z}_{0}}\left(\overline{z}_{0}\right)f_{\overline{W}_{0}}\left(\overline{w}_{0}\right)$ since $\overline{Z}_{0}$ and $\overline{W}_{0}$ are independent under our model. Note these values are not equal to their counterparts in the parameters known case.

4.5 Robustness of the EWMA $\overline{X}$ Chart

George Box is known for making the statement that “All models are wrong, some models are useful.” The study of robustness of a statistical model has been reported in the literature in particular that of the Normal model when used with control charts. Burr[7] discussed the robustness of control chart limits and found them to be robust to the Normal model. He suggested that unless the distribution of the quality measurement is extremely non-Normal control limits based on the Normal model may be used. Other authors, including Schilling and Nelson[16],Chan,Hapuarachchi and Macpherson[9]and Yourtone and Zimmer[20], have studied and reported the effect of non Normality on control charts. The non Normal models they employed have been discussed in Chapter 3 along with others. Also, Schilling and Nelson[16] made use of the a mixture of two Normal distributions. Their study indicated that sample of size 4 or 5 were sufficient to insure reasonable of the Normal assumption.Montgomery [14] discusses the robustness of the EWMA $\overline{X}$ chart. He found that the chart is robust to the assumption the quality measurement follows a Normal distribution.

4.6 Conclusion

The exponentially weighted moving average (EWMA) $\overline{X}$ chart has been outlined. The control limits, variable and fixed, for the in-control parameters known and estimated cases were given. Integral equations useful in describing the performance of the chart have been given. A discuss was given about the study of the robustness of control charts including the EWMA $\overline{X}$ chart.

CHAPTER 5

EWMA$^{(K)}$ $\overline{X}$ CHARTS

5.1 INTRODUCTION

Shamma and Shamma[17] introduced the double EWMA $\overline{X}$ chart.Alevizakos,Chatterjee and Koukouvinos[5] introduced the triple EWMA $\overline{X}$ chart and Alevizakos,Chatterjee and Koukouvinos[6] introduced the quadruple EWMA $\overline{X}$ chart. In the next section, we given a general description of the EWMA$^{(k)}$ $\overline{X}$ chart in which the smoothing parameters $(r_1, \ldots, r_k)$ are not necessarily equal. The aforementions EWMA$^{(k)}$ $\overline{X}$ charts are special cases in which $r_1 = \ldots = r_k$ for $k = 2, 3, 4$.

The EWMA$^{(k)}$ $\overline{X}$ chart is discussed in Section 3. According to Shamma and Shamma[17], the double EWMA $\overline{X}$ chart (EWMA$^{(2)}$ $\overline{X}$ chart) has better out-of-control average run lengths (ARL) for small to moderate shifts in the process mean than the EWMA $\overline{X}$ chart for $r_1 = r_2$. The double EWMA $\overline{X}$ chart (or EWMA$^{(2)}$ $\overline{X}$ chart) is discussed in Section 4.Mahmoud[13] concluded that "the DEWMA (λ) performs better than the EWMA (λ) chart only when a large value of λ is used to design both charts and the process means shifts are smaller than the process standard deviation." Here λ is the smoothing parameter for the chart. Rigdon using simulations showed that the the optimal double EWMA $\overline{X}$ chart is the optimal EWMA $\overline{X}$ chart.

Alevizakos,Chatterjee and Koukouvinos[5] argue the triple EWMA $\overline{X}$ chart has better out-of-control ARL performance for small to moderate shifts in the mean than the double EWMA $\overline{X}$ chart for $r_1 = r_2 = r_3$. The triple EWMA $\overline{X}$ chart (or EWMA$^{(3)}$ $\overline{X}$ chart) is discussed in Section 5. Alevizakos,Chatterjee and Koukouvinos[6] argue the quadruple EWMA $\overline{X}$ chart (or EWMA$^{(4)}$ $\overline{X}$ chart) has a better out-of-control ARL performance for small to moderate shifts in the mean than the triple EWMA $\overline{X}$ chart for $r_1 = r_2 = r_3 = r_4$. The EWMA$^{(4)}$ $\overline{X}$ chart is discussed in Section 6.

Methods are given in Section 7 for designing the EWMA $\overline{X}$ chart that has better out-of-control ARL performance for small to moderate shifts in the mean than a given EWMA$^{(k)}$ $\overline{X}$ chart. We will discuss how the results obtained by Rigdon for the double EWMA $\overline{X}$ chart can be extended to the triple EWMA $\overline{X}$ chart. Some conclusions are drawn in Section 8.

5.2 EWMA$^{(k)}$ $\overline{X}$ Chart with Known Parameters

The EWMA$^{(k)}$ $\overline{X}$ chart with known in-control process parameters is a plot of the points with coordinates $\left(t, U_t^{(k,r_1,\ldots,r_k)}\right)$ for $t = 1, 2, 3, \ldots$, where

$$\begin{aligned}
U_0^{(1,r_1)} &= \mu_0, U_t^{(1,r_1)} = (1 - r_1)\, U_{t-1}^{(1,r_1)} + r_1 \overline{X}_t; \\
U_0^{(2,r_1,r_2)} &= \mu_0, U_t^{(2,r_1,r_2)} = (1 - r_2)\, U_{t-1}^{(2,r_1,r_2)} + r_2 U_t^{(1,r_1)}; \\
&\vdots \\
U_0^{(k,r_1,\ldots,r_k)} &= \mu_0, U_t^{(k,r_1,\ldots,r_k)} = (1 - r_k)\, U_{t-1}^{(k,r_1,\ldots,r_k)} + r_k U_t^{(k-1,r_1,\ldots,r_{k-1})}.
\end{aligned}$$

The chart signals at the first sampling stage t in which $U_t^{(k)} \leq LCL_t^{(k)}$ or $U_t^{(k)} \geq UCL_t^{(k)}$, where

$$\begin{aligned}
LCL_t^{(k,r_1,\ldots,r_k)} &= E\left(U_t^{(k,r_1,\ldots,r_k)}\right) - h\sqrt{V\left(U_t^{(k,r_1,\ldots,r_k)}\right)} \text{ and} \\
UCL_t^{(k,r_1,\ldots,r_k)} &= E\left(U_t^{(k,r_1,\ldots,r_k)}\right) + h\sqrt{V\left(U_t^{(k,r_1,\ldots,r_k)}\right)}
\end{aligned}$$

with the expectation and variance determined under the assumption the process is in-control. Typically these limits depend on t. The fixed limits for these charts are

$$LCL_t^{(k,r_1,\ldots,r_k)} = E\left(U_t^{(k,r_1,\ldots,r_k)}\right) - h\sqrt{\lim_{t\to\infty} V\left(U_t^{(k,r_1,\ldots,r_k)}\right)} \text{ and}$$
$$UCL_t^{(k,r_1,\ldots,r_k)} = E\left(U_t^{(k,r_1,\ldots,r_k)}\right) + h\sqrt{\lim_{t\to\infty} V\left(U_t^{(k,r_1,\ldots,r_k)}\right)}.$$

The chart parameters $r_1, \ldots, r_k$ are called smoothing parameters with $0 < r_i \leq 1$ for $i = 1, \ldots, k$. The chart parameter h is a positive real number. Setting $k = 2$ results in the double EWMA $\overline{X}$ chart of Shamma and Shamma[17] with $r_1 = r_2$. Setting $k = 3$, we have the triple EWMA $\overline{X}$ chart of Alevizakos,Chatterjee and Koukouvinos[5] with $r_1 = r_2 = r_3$, and $k = 4$ the quadruple EWMA $\overline{X}$ chart of Alevizakos,Chatterjee and Koukouvinos[6] with $r_1 = r_2 = r_3 = r_4$. These authors have demonstrated via simulation that EWMA$^{(k)}$ $\overline{X}$ chart outperforms the EWMA $\overline{X}$ chart with smoothing parameter $0 < r_1 \leq 1$.

The following theorems give some general results for the EWMA$^{(k)}$ $\overline{X}$ chart. The proof of these theorems makes use of the Axiom of Induction and the Principle of Double Induction.

Theorem 5.2.1. For $k \geq 1$and $t \geq 1$,

$$U_0^{(k,r_1,\ldots,r_k)} = \mu_0 \text{ and}$$
$$U_t^{(k,r_1,\ldots,r_k)} = (1-r_k)^t \mu_0 + r_k \sum\nolimits_{i=1}^{t} (1-r_k)^{t-i} U_i^{(k-1,r_1,\ldots,r_{k-1})},$$

where $U_i^{(0)} = \overline{X}_i$.

Proof of Theorem 5.2.1. For $k = 1$ and $t = 1$, we have

$$U_1^{(1,r_1)} = (1-r_1)^1 \mu_0 + r_1 \sum\nolimits_{i=1}^{1} (1-r_1)^{1-i} U_i^{(1-1)}.$$

Hence, the theorem is true for $k = 1$ and $t = 1$. Suppose the theorem is true for any arbtrary $k > 1$ and $t = 1$. It follows that

$$U_1^{(k+1,r_1,\ldots,r_{k+1})} = (1-r_{k+1})^1 \mu_0 + r_{k+1} \sum\nolimits_{i=1}^{1} (1-r_{k+1})^{1-i} U_i^{(k,r_1,\ldots,r_k)}.$$

Thus, the theorem is true for $k+1$ and $t = 1$. Suppose the theorem is true for any arbtrary $k > 1$ and arbtrary $t > 1$. We have

$$\begin{aligned} U_{t+1}^{(k,r_1,\ldots,r_k)} &= (1-r_k) U_t^{(k,r_1,\ldots,r_k)} + r_k U_{t+1}^{(k-1,r_1,\ldots,r_{k-1})} \\ &= (1-r_k) \begin{bmatrix} (1-r_k)^t \mu_0 \\ +r_k \sum_{i=1}^{t} (1-r_k)^{t-i} U_i^{(k-1,r_1,\ldots,r_{k-1})} \end{bmatrix} \\ &+ r_k U_{t+1}^{(k-1,r_1,\ldots,r_{k-1})} \\ &= (1-r_k)^{t+1} \mu_0 + r_k \sum\nolimits_{i=1}^{t} (1-r_k)^{t+1-i} U_i^{(k-1,r_1,\ldots,r_{k-1})} \\ &+ r_k U_{t+1}^{(k-1,r_1,\ldots,r_{k-1})} \\ &= (1-r_k)^{t+1} \mu_0 + r_k \sum\nolimits_{i=1}^{t+1} (1-r_k)^{t+1-i} U_i^{(k-1,r_1,\ldots,r_{k-1})}. \end{aligned}$$

Hence, the theorem is true for any arbitrary $k > 1$ and $t+1$. Therefore, by the Principle of Double Induction the theorem holds for all positive integers $k \geq 1$ and positive integers $t \geq 1$.

Theorem 5.2.2. Suppose $r_1 = \ldots = r_{k-1} = 1$, then for $k \geq 2$

$$U_0^{(k,1,\ldots,1,r_k)} = \mu_0, U_t^{(k,1,\ldots,1,r_k)} = (1-r_k) U_{t-1}^{(k,1,\ldots,1,r_k)} + r_k \overline{X}_t.$$

Hence, it follows that EWMA$^{(k,1,\ldots,1,r_k)}$ $\overline{X}$ chart is an EWMA $\overline{X}$ chart with smoothing parameter r_k.

Proof of Theorem 5.2.2. By definition of $U_0^{(k,1,\ldots,1,r_k)}$, we have $U_0^{(k,1,\ldots,1,r_k)} = \mu_0$. We have shown in Theorem 4.3.2 the theorem is true for $k = 1$. Suppose the theorem is true for an arbtrary $k > 1$. We have

$$\begin{aligned} U_1^{(k+1,1,\ldots,1,r_{k+1})} &= (1 - r_{k+1})\, U_0^{(k+1,1,\ldots,1,r_{k+1})} + r_{k+1} U_1^{(k,1,\ldots,1,r_k)} \\ &= (1 - r_{k+1})\, \mu_0 + r_{k+1} U_1^{(k,1,\ldots,1,r_k)} \\ &= (1 - r_{k+1})\, \mu_0 + r_{k+1} \overline{X}_1. \end{aligned}$$

Hence, the theorem is true for $k+1$. Suppose the theorem is true for an arbtrary positive integer $k > 1$ and arbtrary positive integer $t > 1$. We have

$$U_{t+1}^{(k,1,\ldots,1,r_k)} = (1 - r_k)\, U_t^{(k,1,\ldots,1,r_k)} + r_k \overline{X}_{t+1}.$$

Hence, the theorem is true for an arbtrary $k > 1$ and $t + 1$. Therefore, by the Principle of Double Induction the theorem is true for all positive integers k and positive integers t.

To conveniently study the EWMA$^{(k)}$ $\overline{X}$ chart, we introduce a standardized version of the in-control parameters known chart by defining $U_t^{(i)*}$ as

$$U_t^{(i)*} = \frac{U_t^{(i)} - \mu_0}{\sigma_0/\sqrt{n}},$$

for $i = 2, 3, \ldots, k$. It follows that

$$U_0^{(i)*} = 0 \text{ and } U_t^{(i)*} = (1 - r_i)\, U_{t-1}^{(i)*} + r_i U_t^{(i-1)*},$$

with $U_i^{(1)*} = Z_i + \delta$, where

$$Z_t = \frac{\overline{X}_t - \mu}{\sigma_0/\sqrt{n}} \sim N(0, 1) \text{ and } \delta = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}.$$

The process is in a state of statistical in control if $\delta = 0$. We will refer to this chart as the EWMA$^{(k)*}$ Z chart. The EWMA$^{(k)*}$ Z chart with known in-control process parameters is a plot of the points with coordinates $\left(t, U_t^{(k)*}\right)$ for $t = 1, 2, 3, \ldots$, where

$$\begin{aligned}
&U_0^{(1)*} = 0, U_t^{(1)*} = (1 - r_1)\, U_{t-1}^{(1)*} + r_1 (Z_t + \delta)\,; \\
&U_0^{(2)*} = 0, U_t^{(2)*} = (1 - r_2)\, U_{t-1}^{(2)*} + r_2 U_t^{(1)*}; \\
&\quad\vdots \\
&U_0^{(k)*} = 0, U_t^{(k)*} = (1 - r_k)\, U_{t-1}^{(k)*} + r_k U_t^{(k-1)*}.
\end{aligned}$$

The chart signals at the first sampling stage t in which $U_t^{(k)*} \leq LCL_t^{(k)*}$ or $U_t^{(k)*} \geq UCL_t^{(k)*}$, where the variable control limits are

$$LCL_t^{(k)*} = -h\sqrt{V\left(U_t^{(k)*}\right)} \text{ and } UCL_t^{(k)*} = h\sqrt{V\left(U_t^{(k)*}\right)}$$

with the variance determined under the assumption the process is in-control. The fixed control limits are

$$LCL^{(k)*} = -h\sqrt{\lim_{t\to\infty} V\left(U_t^{(k)*}\right)} \text{ and } UCL^{(k)*} = h\sqrt{\lim_{t\to\infty} V\left(U_t^{(k)*}\right)}.$$

The chart smoothing parameters $r_1, \ldots, r_k$ and the chart parameter h are the same as the chart parameters for the EWMA$^{(k)}$ $\overline{X}$ chart. The unobservable EWMA$^{(k)*}$ Z chart is equivalent to the EWMA$^{(k)}$ $\overline{X}$ chart in the sense that at time t both charts signal or both charts do not signal.

In the parameters estimated case, we define for $i = 1, 2, 3, \ldots$ and $t = 1, 2, 3, \ldots$

$$U_t^{(i)*} = c_{m,n} \frac{U_t^{(i)} - \overline{\overline{X}}_0}{\overline{V}_0/\sqrt{n}},$$

where

$$c_{m,n} = \frac{\sqrt{2}\Gamma\left(\frac{m(n-1)}{2}\right)}{\sqrt{m(n-1)}\Gamma\left(\frac{m(n-1)-1}{2}\right)}.$$

The value $c_{m,n}$ is what one might refer to as a unbiasing contant in the sense that

$$E\left(U_t^{(i)*}\middle|U_t^{(i)}\right) = \frac{U_t^{(i)} - \mu_0}{\sigma_0/\sqrt{n}}.$$

The variable control limits for this chart are

$$LCL_t^{(k)*} = -h\sqrt{V\left(U_t^{(k)*}\right)} \text{ and } UCL_t^{(k)*} = h\sqrt{V\left(U_t^{(k)*}\right)}$$

with the variance determined under the assumption the process is in-control. The fixed control limits are

$$LCL^{(k)*} = -h\sqrt{\lim_{t\to\infty} V\left(U_t^{(k)*}\right)} \text{ and } UCL^{(k)*} = h\sqrt{\lim_{t\to\infty} V\left(U_t^{(k)*}\right)}.$$

5.3 EWMA$^{(2)}$ or Double EWMA $\overline{X}$ Chart

Shamma and Shamma[17] introduced the double EWMA chart with equal smoothing parameters. We extend their definition of a double EWMA chart to unequal smoothing parameters. The double EWMA $\overline{X}$ chart (or EWMA$^{(2)}$ $\overline{X}$ chart) with known in-control process parameters is a plot of the points with coordinates $\left(t, U_t^{(2,r_1,r_2)}\right)$ for $t = 1, 2, 3, \ldots$, where

$$U_0^{(1,r_1)} = \mu_0, U_t^{(1,r_1)} = (1 - r_1)\, U_{t-1}^{(1,r_1)} + r_1\overline{X}_t;$$
$$U_0^{(2,r_1,r_2)} = \mu_0, U_t^{(2,r_1,r_2)} = (1 - r_2)\, U_{t-1}^{(2,r_1,r_2)} + r_2 U_t^{(1,r_1)}.$$

The chart signals at the first sampling stage t in which $U_t^{(2)} \leq LCL_t^{(2)}$ or $U_t^{(2)} \geq UCL_t^{(2)}$, where

$$LCL_t^{(2,r_1,r_2)} = E\left(U_t^{(2,r_1,r_2)}\right) - h\sqrt{V\left(U_t^{(2,r_1,r_2)}\right)} \text{ and}$$
$$UCL_t^{(2)} = E\left(U_t^{(2,r_1,r_2)}\right) + h\sqrt{V\left(U_t^{(2,r_1,r_2)}\right)}$$

with the expectation and variance determined under the assumption the process is in-control. The following theorems are useful in analyzing the chart.

Theorem 5.3.1. $U_t^{(2,r_1,r_2)}$ and $U_t^{(2,r_1,r_2)*}$ can be expressed as

$$U_t^{(2,r_1,r_2)} = \begin{cases} \left[(1-r_1) r_2 B_{t-1}^{(2)} + (1-r_2)^t\right] \mu_0 \\ +r_1 r_2 \sum_{i=1}^t B_{t-i}^{(2)} \overline{X}_i, \text{ if } r_1 \neq r_2; \text{ and} \\ (1+tr_2)(1-r_2)^t \mu_0 \\ +r_2^2 \sum_{i=1}^t (t+1-i)(1-r_2)^{t-i} \overline{X}_i, \\ \text{if } r_1 = r_2; \end{cases}$$

$$U_t^{(2,r_1,r_2)*} = \begin{cases} r_1 r_2 \sum_{i=1}^t B_{t-i}^{(2)} (Z_i+\delta), \text{ if } r_1 \neq r_2; \\ r_2^2 \sum_{i=1}^t (t+1-i)(1-r_2)^{t-i} (Z_i+\delta), \\ \text{if } r_1 = r_2; \end{cases},$$

where

$$B_{t-i}^{(2)} = \sum_{j=0}^{t-i} (1-r_1)^{t-i-j} (1-r_2)^j$$

with $B_{t-i}^{(2)} = 1$ for $i = t$ for $t = 1, 2, 3, \ldots$.

Proof of Theorem 5.3.1. For $t = 1$ and $r_1 \neq r_2$, we have

$$\begin{aligned} U_1^{(2,r_1,r_2)} &= (1-r_2)\mu_0 + r_2 U_1^{(1,r_1)} = (1-r_2)\mu_0 + r_2\left[(1-r_1)\mu_0 + r_1\overline{X}_1\right] \\ &= \left[(1-r_1) r_2 + (1-r_2)\right]\mu_0 + r_1 r_2 \overline{X}_1 \\ &= \left[(1-r_1) r_2 B_{1-1}^{(2)} + (1-r_2)\right]\mu_0 + r_1 r_2 \sum_{i=1}^1 B_{1-i}^{(2)} \overline{X}_i. \end{aligned}$$

Hence, the theorem is true for $t = 1$ and $r_1 \neq r_2$. Suppose the theorem is true for an arbitrary $t > 1$ and $r_1 \neq r_2$. It follows that

$$\begin{aligned}
U_{t+1}^{(2,r_1,r_2)} &= (1-r_2)\,U_t^{(2,r_1,r_2)} + r_2 U_{t+1}^{(1,r_1)} \\
&= (1-r_2)\left[\begin{array}{c} \left[(1-r_1)\,r_2 B_{t-1}^{(2)} + (1-r_2)^t\right]\mu_0 \\ +r_1 r_2 \sum_{i=1}^{t} B_{t-i}^{(2)} \overline{X}_i \end{array}\right] \\
&+ r_2\left[(1-r_1)^{t+1}\mu_0 + r_1 \sum_{i=1}^{t+1} (1-r_1)^{t+1-i}\,\overline{X}_i\right] \\
&= \left[(1-r_1)\,r_2\left[(1-r_1)^t + (1-r_2)\,B_{t-1}^{(2)}\right] + (1-r_2)^{t+1}\right]\mu_0 \\
&+ r_1 r_2\left[\sum_{i=1}^{t}\left[B_{t-i}^{(2)}\,(1-r_2) + (1-r_1)^{t+1-i}\right]\overline{X}_i + \overline{X}_{t+1}\right].
\end{aligned}$$

Observe that

$$\begin{aligned}
&(1-r_1)^t + (1-r_2)\,B_{t-1}^{(2)} \\
&= (1-r_1)^t\,(1-r_2)^0 + \sum_{j=0}^{t-1} (1-r_1)^{t-1-j}\,(1-r_2)^{j+1} \\
&= (1-r_1)^t\,(1-r_2)^0 + (1-r_1)^{t-1}\,(1-r_2)^1 \\
&+ \ldots + (1-r_1)^0\,(1-r_2)^t \\
&= \sum_{j=0}^{t+1-1} (1-r_1)^{t+1-1-j}\,(1-r_2)^j = B_{t+1-1}^{(2)}.
\end{aligned}$$

Further observe that

$$\begin{aligned}
&B_{t-i}^{(2)}\,(1-r_2) + (1-r_1)^{t+1-i} \\
&= \sum_{j=0}^{t-i} (1-r_1)^{t-i-j}\,(1-r_2)^{j+1} + (1-r_1)^{t+1-i} \\
&= (1-r_1)^{t+1-i}\,(1-r_2)^0 + (1-r_1)^{t-i}\,(1-r_2)^1 \\
&+ (1-r_1)^{t-i-1}\,(1-r_2)^2 + \ldots + (1-r_1)^0\,(1-r_2)^{t-i+1} \\
&= \sum_{j=0}^{t+1-i} (1-r_1)^{t+1-i-j}\,(1-r_2)^j = B_{t+1-i}^{(2)}.
\end{aligned}$$

Therefore,

$$U_{t+1}^{(2,r_1,r_2)} = \left[(1-r_1)\,r_2 B_{t+1-1}^{(2)} + (1-r_2)^{t+1}\right]\mu_0 + r_1 r_2 \sum_{i=1}^{t+1} B_{t+1-i}^{(2)} \overline{X}_i.$$

It follows by the Axiom of Induction that the theorem is true for all positive integers $t \geq 1$.

For the case in which $t = 1$ and $r_1 = r_2$, we have

$$\begin{aligned} U_1^{(2,r_2,r_2)} &= (1 - r_2)\,\mu_0 + r_2 U_1^{(1,r_2)} \\ &= (1 - r_2)\,\mu_0 + r_2\left[(1 - r_2)\,\mu_0 + r_2 \overline{X}_1\right] \\ &= (1 + 1r_2)\,(1 - r_2)\,\mu_0 \\ &+ r_2^2 \sum\nolimits_{i=1}^{1} (1 + 1 - i)\,(1 - r_2)^{1-i}\,\overline{X}_i \\ &= [(1 - r_1)\,r_2 + (1 - r_2)]\,\mu_0 + r_1 r_2 \overline{X}_1. \end{aligned}$$

Hence the theorem is true for $t = 1$ and $r_1 = r_2$. Suppose the theorem is true for an arbtrary $t > 1$. We have

$$\begin{aligned} U_{t+1}^{(2,r_2,r_2)} &= (1 - r_2)\,U_t^{(2,r_2,r_2)} + r_2 U_{t+1}^{(1,r_2)} \\ &= (1 + t r_2)\,(1 - r_2)^{t+1}\,\mu_0 \\ &+ r_2^2 \sum\nolimits_{i=1}^{t} (t + 1 - i)\,(1 - r_2)^{t+1-i}\,\overline{X}_i \\ &+ r_2\left[(1 - r_2)^{t+1}\,\mu_0 + r_2 \sum\nolimits_{i=1}^{t+1} (1 - r_2)^{t+1-i}\,\overline{X}_i\right] \\ &= (1 + (t + 1)\,r_2)\,(1 - r_2)^{t+1}\,\mu_0 \\ &+ r_2^2 \sum\nolimits_{i=1}^{t+1} ((t + 1) + 1 - i)\,(1 - r_2)^{t+1-i}\,\overline{X}_i. \end{aligned}$$

Thus, the theorem is true for $t + 1$. Therefore, by the Axiom of Induction, the theorem is true for all positive integers t.

The variance of $U_t^{(2,r_1,r_2)}$ for $r_1 \neq r_2$ is determined as follows.

$$V\left(U_t^{(2,r_1,r_2)}\right) = r_1^2 r_2^2 \sum\nolimits_{i=1}^{t} \left(B_{t-i}^{(2)}\right)^2 \frac{\sigma^2}{n}$$

Observe that we can write

$$
\begin{aligned}
B_{t-i}^{(2)} &= \sum\nolimits_{j=0}^{t-i} (1-r_1)^{t-i-j} (1-r_2)^j \\
&= (1-r_1)^{t-i} \sum\nolimits_{j=0}^{t-i} \left(\frac{1-r_2}{1-r_1}\right)^j \\
&= (1-r_1)^{t-i} \frac{\left(\frac{1-r_2}{1-r_1}\right) - \left(\frac{1-r_2}{1-r_1}\right)^{t-i+1}}{1-\frac{1-r_2}{1-r_1}} \\
&= (1-r_1)^{t-i} \frac{1-\left(\frac{1-r_2}{1-r_1}\right)^{t-i}}{\frac{1-r_1}{1-r_2}-1} \\
&= \frac{(1-r_2)}{r_2-r_1}\left[(1-r_1)^{t-i} - (1-r_2)^{t-i}\right].
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\left(B_{t-i}^{(2)}\right)^2 &= \frac{(1-r_2)^2}{(r_2-r_1)^2}\left[(1-r_1)^{t-i} - (1-r_2)^{t-i}\right]^2 \\
&= \frac{(1-r_2)^2}{(r_2-r_1)^2}\left[\begin{array}{c} (1-r_1)^{2(t-i)} + (1-r_2)^{2(t-i)} \\ -2(1-r_1)^{t-i}(1-r_2)^{t-i} \end{array}\right].
\end{aligned}
$$

Using these results, we have that

$$
V\left(U_t^{(2,r_1,r_2)}\right) = r_1^2 r_2^2 \frac{(1-r_2)^2}{(r_2-r_1)^2}\left[\begin{array}{c} \sum_{i=1}^t (1-r_1)^{2(t-i)} + \sum_{i=1}^t (1-r_2)^{2(t-i)} \\ -2\sum_{i=1}^t (1-r_1)^{t-i}(1-r_2)^{t-i} \end{array}\right]\frac{\sigma^2}{n}.
$$

Observe that

$$\begin{aligned}
&\sum\nolimits_{i=1}^{t}(1-r_1)^{2(t-i)} \\
&=(1-r_1)^{2t}\sum\nolimits_{i=1}^{t}\left[(1-r_1)^{-2}\right]^i \\
&=(1-r_1)^{2t}\frac{(1-r_1)^{-2}-(1-r_1)^{-2(t+1)}}{1-(1-r_1)^{-2}} \\
&=(1-r_1)^{2t}\frac{1-(1-r_1)^{-2t}}{(1-r_1)^2-1} \\
&=\frac{(1-r_1)^{2t}-1}{(1-r_1)^2-1}=\frac{1-(1-r_1)^{2t}}{1-(1-r_1)^2}.
\end{aligned}$$

Similarly, we have

$$\sum\nolimits_{i=1}^{t}(1-r_2)^{2(t-i)}=\frac{1-(1-r_2)^{2t}}{1-(1-r_2)^2}.$$

Next observe that

$$\begin{aligned}
&\sum\nolimits_{i=1}^{t}(1-r_1)^{t-i}(1-r_2)^{t-i} \\
&=(1-r_1)^t(1-r_2)^t\sum\nolimits_{i=1}^{t}\left([(1-r_1)(1-r_2)]^{-1}\right)^i \\
&=(1-r_1)^t(1-r_2)^t\frac{[(1-r_1)(1-r_2)]^{-1}-[(1-r_1)(1-r_2)]^{-(t+1)}}{1-[(1-r_1)(1-r_2)]^{-1}} \\
&=(1-r_1)^t(1-r_2)^t\frac{1-[(1-r_1)(1-r_2)]^{-t}}{[(1-r_1)(1-r_2)]-1} \\
&=\frac{1-[(1-r_1)(1-r_2)]^t}{1-[(1-r_1)(1-r_2)]}.
\end{aligned}$$

It follows that

$$V\left(U_t^{(2,r_1,r_2)}\right)=r_1^2r_2^2\frac{(1-r_2)^2}{(r_2-r_1)^2}\left[\begin{array}{c}\frac{1-(1-r_1)^{2t}}{1-(1-r_1)^2}+\frac{1-(1-r_2)^{2t}}{1-(1-r_2)^2} \\ -2\frac{1-(1-r_1)^t(1-r_2)^t}{1-(1-r_1)(1-r_2)}\end{array}\right]\frac{\sigma^2}{n}.$$

Theorem 5.3.2. The variance of $U_t^{(2,r_1,r_2)}$ is

$$V\left(U_t^{(2,r_1,r_2)}\right) = \begin{cases} \frac{r_1^2 r_2^2}{(r_2-r_1)^2} \left[\begin{array}{c} \frac{(1-r_1)^2\left[1-(1-r_1)^{2t}\right]}{1-(1-r_1)^2} + \frac{(1-r_2)^2\left[1-(1-r_2)^{2t}\right]}{1-(1-r_2)^2} \\ -2\frac{(1-r_1)(1-r_2)\left[1-(1-r_1)^t(1-r_2)^t\right]}{1-(1-r_1)(1-r_2)} \end{array} \right] \frac{\sigma^2}{n}, \\ \text{if } r_1 \neq r_2; \\ \left(\frac{r_2}{(2-r_2)^3} \left[\begin{array}{c} \left(2r_2 + 2r_2^2 t\right)(1-r_2)^{2t} \\ +2 - \left(\begin{array}{c} 4r_2^2t^2 - 4r_2^3t^2 + r_2^4t^2 \\ +4r_2t + r_2^2 \end{array} \right) \\ \times (1-r_2)^{2t} \\ +r_2^2 - 2r_2 + 2(1-r_2)^{2t} \end{array} \right] \right) \frac{\sigma^2}{n}, \\ \text{if } r_1 = r_2; \end{cases}$$

for $t = 1, 2, 3, \ldots$ and

$$\lim_{t\to\infty} V\left(U_t^{(2,r_1,r_2)}\right) = \begin{cases} \frac{r_1^2 r_2^2}{(r_2-r_1)^2} \left(\frac{(1-r_1)^2}{1-(1-r_1)^2} + \frac{(1-r_2)^2}{1-(1-r_2)^2} - 2\frac{(1-r_1)(1-r_2)}{1-(1-r_1)(1-r_2)} \right) \frac{\sigma^2}{n}, \\ \text{if } r_1 \neq r_2; \\ \frac{r_2\left(2-2r_2+r_2^2\right)}{(2-r_2)^3} \frac{\sigma^2}{n}, \\ \text{if } r_1 = r_2. \end{cases}$$

Also,

$$V\left(U_t^{(2,r_1,r_2)*}\right) = \frac{V\left(U_t^{(2)}\right)}{\sigma^2/n} \text{ and } \lim_{t\to\infty} V\left(U_t^{(2)*}\right) = \frac{\lim_{t\to\infty} V\left(U_t^{(2)}\right)}{\sigma^2/n}.$$

$$\frac{r_2\left(2-2r_2+r_2^2\right)}{(2-r_2)^3} = \frac{r_2\left[1+\left(1-2r_2+r_2^2\right)\right]}{(2-r_2)^3}$$

It follow that (1) the variable and (2) fixed control limits for the double EWMA $\overline{X}$ chart are

$$(1)\ LCL_t^{(2,r_1,r_2)} = \mu_0 - h\sqrt{V\left(U_t^{(2,r_1,r_2)}\right)} \text{ and}$$
$$UCL_t^{(2,r_1,r_2)} = \mu_0 + h\sqrt{V\left(U_t^{(2,r_1,r_2)}\right)};$$
$$(2)\ LCL_t^{(2,r_1,r_2)} = \mu_0 - h\sqrt{\lim_{t\to\infty} V\left(U_t^{(2,r_1,r_2)}\right)} \text{ and}$$
$$UCL_t^{(2,r_1,r_2)} = \mu_0 + h\sqrt{\lim_{t\to\infty} V\left(U_t^{(2,r_1,r_2)}\right)},$$

where the variance is determined when the process is in a state of statistical in control. The (3) variable and (4) fixed control limits for the double EWMA$^{(2)}$ Z chart are

$$(3)\ LCL_t^{(2,r_1,r_2)*} = -h\sqrt{V\left(U_t^{(2,r_1,r_2)*}\right)} \text{ and } UCL_t^{(2)*} = h\sqrt{V\left(U_t^{(2,r_1,r_2)*}\right)} \text{and}$$
$$(4)\ LCL_t^{(2,r_1,r_2)*} = -h\sqrt{\lim_{t\to\infty} V\left(U_t^{(2,r_1,r_2)*}\right)} \text{ and } UCL_t^{(2)*} = h\sqrt{\lim_{t\to\infty} V\left(U_t^{(2,r_1,r_2)*}\right)},$$

where the variance is determined when the process is in-control. These results are presented in Shamma and Shamma.

5.4 TRIPLE EWMA $\overline{X}$ CHART

Alevizakos,Chatterjee and Koukouvinos[5] introduced the triple EWMA $\overline{X}$ chart. The triple EWMA $\overline{X}$ chart (or EWMA$^{(3)}$ $\overline{X}$ chart) with known in-control process parameters is a plot of the points with coordinates $\left(t, U_t^{(3,r_1,r_2,r_3)}\right)$ for $t = 1, 2, 3, \ldots$, where

$$U_0^{(1,r_1)} = \mu_0,\ U_t^{(1,r_1)} = (1 - r_1)\, U_{t-1}^{(1,r_1)} + r_1\overline{X}_t;$$
$$U_0^{(2,r_1,r_2)} = \mu_0,\ U_t^{(2,r_1,r_2)} = (1 - r_2)\, U_{t-1}^{(2,r_1,r_2)} + r_2 U_t^{(1,r_1)};$$
$$U_0^{(3,r_1,r_2,r_3)} = \mu_0,\ U_t^{(3,r_1,r_2,r_3)} = (1 - r_3)\, U_{t-1}^{(3,r_1,r_2,r_3)} + r_3 U_t^{(2,r_1,r_2)}.$$

The chart signals at the first sampling stage t in which $U_t^{(3)} \leq LCL_t^{(3)}$ or $U_t^{(3)} \geq UCL_t^{(3)}$, where the variable limits are

$$LCL_t^{(3,r_1,r_2,r_3)} = E\left(U_t^{(3,r_1,r_2,r_3)}\right) - h\sqrt{V\left(U_t^{(3,r_1,r_2,r_3)}\right)};$$
$$UCL_t^{(3,r_1,r_2,r_3)} = E\left(U_t^{(3,r_1,r_2,r_3)}\right) + h\sqrt{V\left(U_t^{(3,r_1,r_2,r_3)}\right)}$$

and fixed limits

$$LCL_t^{(3,r_1,r_2,r_3)} = E\left(U_t^{(3,r_1,r_2,r_3)}\right) - h\sqrt{\lim_{t\to\infty} V\left(U_t^{(3,r_1,r_2,r_3)}\right)};$$
$$UCL_t^{(3,r_1,r_2,r_3)} = E\left(U_t^{(3,r_1,r_2,r_3)}\right) + h\sqrt{\lim_{t\to\infty} V\left(U_t^{(3,r_1,r_2,r_3)}\right)}$$

with the expectation and variance determined under the assumption the process is in-control.

There are five possible senerios for the smoothing parameters r_1, r_2, and r_3. They are

$$\begin{bmatrix} (1) & r_1 \neq r_2 & r_1 \neq r_3 & r_2 \neq r_3 \\ (2) & r_1 \neq r_2 & r_1 \neq r_3 & r_2 = r_3 \\ (3) & r_1 \neq r_2 & r_1 = r_3 & r_2 \neq r_3 \\ (4) & r_1 = r_2 & r_1 \neq r_3 & r_2 \neq r_3 \\ (5) & r_1 = r_2 & r_1 = r_3 & r_2 = r_3 \end{bmatrix}.$$

We will only examine cases (1) and (5) leaving the others for future research. We plan to use a simulation method similar to the one used by Rigdon to investigate our conjecture that the optimal triple EWMA $\overline{X}$ chart is the optimal EWMA $\overline{X}$ chart.

The variance of $U_t^{(3,r_1,r_2,r_3)}$ can be determined by

$$\lim_{t\to\infty} V\left(U_t^{(3,r_1,r_2,r_3)}\right) = \begin{cases} r_1^2 r_2^2 r_3^2 [\frac{(1-r_1)^4}{(r_1-r_2)^2(r_1-r_3)^2} \\ +\frac{(1-r_2)^4}{(r_1-r_2)^2(r_3-r_2)^2} \\ +\frac{(1-r_3)^4}{(r_1-r_3)^2(r_3-r_2)^2\left[1-(1-r_3)^2\right]} \\ -2\frac{(1-r_1)^2(1-r_2)^2}{(r_1-r_2)^2(r_1-r_3)(r_3-r_2)\left[1-(1-r_1)^2(1-r_2)^2\right]} \\ -2\frac{(1-r_1)^2(1-r_2)^2}{(r_1-r_2)^2(r_1-r_3)(r_3-r_2)\left[1-(1-r_1)^2(1-r_2)^2\right]} \\ +2\frac{(1-r_1)^2(1-r_3)^2}{(r_1-r_2)(r_1-r_3)^2(r_3-r_2)\left[1-(1-r_1)^2(1-r_2)^2\right]}]\frac{\sigma^2}{n} \\ \text{if } r_1 \neq r_2;\, r_1 \neq r_3; \text{ and } r_2 \neq r_3; \\ r_3^6 \begin{bmatrix} \frac{6(1-r_3)^6 r_3}{(2-r_3)^5} + \frac{12(1-r_3)^4 r_3^2}{(2-r_3)^4} \\ +\frac{7(1-r_3)^2 r_3^3}{(2-r_3)^4} + \frac{r_3^4}{(2-r_3)^{42}} \end{bmatrix} \frac{\sigma^2}{n} \quad . \\ \text{if } r_1 = r_2 = r_3. \end{cases}$$

These results are presented in Alevizakos,Chatterjee and Koukouvinos[5].

The triple EWMA $\overline{X}$ chart (or EWMA$^{(3)}$ $\overline{X}$ chart) with estimated in-control process parameters is a plot of the points with coordinates $\left(t, U_t^{(3)}\right)$ for $t = 1, 2, 3, \ldots$, where

$$U_0^{(1,r_1)} = \overline{\overline{X}}_0,\, U_t^{(1,r_1)} = (1-r_1)\, U_{t-1}^{(1,r_1)} + r_1 \overline{X}_t;$$
$$U_0^{(2,r_1,r_2)} = \overline{\overline{X}}_0,\, U_t^{(2,r_1,r_2)} = (1-r_2)\, U_{t-1}^{(2,r_1,r_2)} + r_2 U_t^{(1,r_1)};$$
$$U_0^{(3,r_1,r_2,r_3)} = \overline{\overline{X}}_0,\, U_t^{(3,r_1,r_2,r_3)} = (1-r_3)\, U_{t-1}^{(3,r_1,r_2,r_3)} + r_3 U_t^{(2,r_1,r_2)}.$$

The chart signals at the first sampling stage t in which $U_t^{(3)} \leq LCL_t^{(3)}$ or $U_t^{(3)} \geq UCL_t^{(3)}$,

where

$$
\begin{aligned}
LCL_t^{(3,r_1,r_2,r_3)} &= E\left(U_t^{(3,r_1,r_2,r_3)}\right) \\
&- h\sqrt{\frac{V\left(U_t^{(3,r_1,r_2,r_3)}\right)}{\sigma^2/n}}\frac{\overline{V}_0^{1/2}/c_{4,m}}{\sqrt{n}} \text{ and} \\
UCL_t^{(3,r_1,r_2,r_3)} &= E\left(U_t^{(3,r_1,r_2,r_3)}\right) \\
&+ h\sqrt{\frac{V\left(U_t^{(3,r_1,r_2,r_3)}\right)}{\sigma^2/n}}\frac{\overline{V}_0^{1/2}/c_{4,m}}{\sqrt{n}}
\end{aligned}
$$

with the expectation and variance determined under the assumption the process is in-control.

5.5 Quadruple EWMA $\overline{X}$ Chart

Alevizakos,Chatterjee and Koukouvinos[6] introduced the quadruple EWMA $\overline{X}$ chart. The quadruple EWMA $\overline{X}$ chart (or EWMA$^{(4)}$ $\overline{X}$ chart) with known in-control process parameters is a plot of the points with coordinates $\left(t, U_t^{(4)}\right)$ for $t = 1, 2, 3, \ldots$, where

$$
\begin{aligned}
U_0^{(1,r_1)} &= \mu_0, U_t^{(1,r_1)} = (1-r_1)\,U_{t-1}^{(1,r_1)} + r_1\overline{X}_t; \\
U_0^{(2,r_1,r_2)} &= \mu_0, U_t^{(2,r_1,r_2)} = (1-r_2)\,U_{t-1}^{(2,r_1,r_2)} \\
&+ r_2U_t^{(1,r_1)}; \\
U_0^{(3,r_1,r_2,r_3)} &= \mu_0, U_t^{(3,r_1,r_2,r_3)} = (1-r_3)\,U_{t-1}^{(3,r_1,r_2,r_3)} \\
&+ r_3U_t^{(2,r_1,r_2)}; \\
U_0^{(4,r_1,r_2,r_3,r_4)} &= \mu_0, U_t^{(4,r_1,r_2,r_3,r_4)} = (1-r_4)\,U_{t-1}^{(4,r_1,r_2,r_3,r_4)} \\
&+ r_4U_t^{(3,r_1,r_2,r_3)}.
\end{aligned}
$$

The variable limits chart signals at the first sampling stage t in which $U_t^{(4,r_1,r_2,r_3,r_4)} \leq LCL_t^{(4,r_1,r_2,r_3,r_4)}$ or $U_t^{(4,r_1,r_2,r_3,r_4)} \geq UCL_t^{(4,r_1,r_2,r_3,r_4)}$, where

$$LCL_t^{(4)} = E\left(U_t^{(4,r_1,r_2,r_3,r_4)}\right) - h\sqrt{V\left(U_t^{(4,r_1,r_2,r_3,r_4)}\right)} \text{ and}$$
$$UCL_t^{(4)} = E\left(U_t^{(4,r_1,r_2,r_3,r_4)}\right) + h\sqrt{V\left(U_t^{(4,r_1,r_2,r_3,r_4)}\right)}$$

with the expectation and variance determined under the assumption the process is in a state of statistical in control. One can show that when the process is in-control $E\left(U_t^{(4,r_1,r_2,r_3,r_4)}\right) = \mu_0$. The fixed limits chart uses the control limits

$$LCL^{(4,r_1,r_2,r_3,r_4)} = \mu_0 - h\sqrt{\lim_{t\to\infty} V\left(U_t^{(4,r_1,r_2,r_3,r_4)}\right)} \text{ and}$$
$$UCL^{(4,r_1,r_2,r_3,r_4)} = \mu_0 + h\sqrt{\lim_{t\to\infty} V\left(U_t^{(4,r_1,r_2,r_3,r_4)}\right)}.$$

If $r_1 = r_2 = r_3 = r_4$, then according to [6]

$$\lim_{t\to\infty} V\left(U_t^{(4)}\right) = \frac{r_4^8}{36}\left[\begin{array}{c} \frac{720d^5}{(1-d)^7} + \frac{2520d^4}{(1-d)^6} + \frac{3312d^3}{(1-d)^5} \\ +\frac{1980d^2}{(1-d)^4} + \frac{50d}{(1-d)^3} + \frac{36}{(1-d)^2} \end{array}\right]\frac{\sigma^2}{n},$$

where $d = (1-r_4)^2$. Observe that

$$\begin{aligned} &\lim_{t\to\infty} V\left(U_t^{(4)}\right) \\ &= \frac{r_4^8}{36}\left[\begin{array}{c} \frac{720(1-r_4)^{10}}{\left(1-(1-r_4)^2\right)^7} + \frac{2520(1-r_4)^8}{\left(1-(1-r_4)^2\right)^6} + \frac{3312(1-r_4)^6}{\left(1-(1-r_4)^2\right)^5} \\ +\frac{1980(1-r_4)^4}{\left(1-(1-r_4)^2\right)^4} + \frac{50(1-r_4)^2}{\left(1-(1-r_4)^2\right)^3} + \frac{36}{\left(1-(1-r_4)^2\right)^2} \end{array}\right]\frac{\sigma^2}{n}. \end{aligned}$$

5.6 Design Methods

We examine three methods for designing the EWMA $\overline{X}$ chart for comparison with the EWMA$^{(k)}$ $\overline{X}$ chart. The first methods sets the smoothing parameter r of the EWMA $\overline{X}$

chart equal to the product of the smoothing parameters of the EWMA$^{(k)}$ $\overline{X}$ chart ($r = r_1 \cdots r_k$). The method can be used whether the variable or fixed control limits are used. A second method chooses r such that the variance the chart statistic U_t of the EWMA $\overline{X}$ chart is equal to the variance of the EWMA$^{(k)}$ $\overline{X}$ chart statistic $U_t^{(k,r_1,\ldots,r_k)}$. This method was studied by Knoth[11]. The value of r for the EWMA $\overline{X}$ chart is the solution to the equation

$$\frac{r}{2-r} = \lim_{t\to\infty} V\left(U_t^{(2,r_k,\ldots,r_k)}\right).$$

It follows that

$$r = \frac{2\lim_{t\to\infty} V\left(U_t^{(2,r_k,\ldots,r_k)}\right)}{1+\lim_{t\to\infty} V\left(U_t^{(2,r_k,\ldots,r_k)}\right)}.$$

The r for the EWMA $\overline{X}$ chart as a function of the smoothing parameters for the double ($r_1 = r_2$), triple ($r_1 = r_2 = r_3$), and quadruple ($r_1 = r_2 = r_3 = r_4$) EWMA $\overline{X}$ charts using methods (1) and (2) are given below .

$$\left[\begin{array}{ccc}
\text{EWMA} & \text{Method 1} & \text{Method 2} \\
\text{Double} & r = r_1 r_2 & r = \frac{2r_2\left(2-2r_2+r_2^2\right)}{(2-r_2)^3+r_2\left(2-2r_2+r_2^2\right)} \\
\text{Triple} & r = r_1 r_2 r_3 & \begin{array}{c} r = \frac{6(1-r_3)^6 r_3^6}{(2-r_3)^5} + \frac{12(1-r_3)^4 r_3^8}{(2-r_3)^4} \\ +\frac{7(1-r_3)^2 r_3^9}{(2-r_3)^4} + \frac{r_3^{10}}{(2-r_3)^{42}} \end{array} \\
\text{Quadruple} & r = r_1 r_2 r_3 r_4 & r = \frac{r_4^8}{36}\left[\begin{array}{c} \frac{720(1-r_4)^{10}}{\left(1-(1-r_4)^2\right)^7} + \frac{2520(1-r_4)^8}{\left(1-(1-r_4)^2\right)^6} \\ +\frac{3312(1-r_4)^6}{\left(1-(1-r_4)^2\right)^5} + \frac{1980(1-r_4)^4}{\left(1-(1-r_4)^2\right)^4} \\ +\frac{50(1-r_4)^2}{\left(1-(1-r_4)^2\right)^3} + \frac{36}{\left(1-(1-r_4)^2\right)^2} \end{array}\right]
\end{array}\right]$$

We first look at some examples for $k = 2$ when $r_1 \neq r_2$. Simulation was used to estimate the ARL's the double EWMA $\overline{X}$ chart with variable control limits having chart parameters $r_1 = 0.2$, $r_2 = 0.4$, and $h^{(2)} = 2.44$ for $\delta = 0.0, 0.1, ..., 1.0$. The in-control

ARL of the chart was estimated to be 198.11. The ARL's of the the EWMA $\overline{X}$ chart with variable control limits having chart parameters $r = r_1 * r_2 = 0.08$ and $h = 2.415$ with estimated in-control ARL of 199.10. For both charts, fifty thousand $(50,000)$ simulations were used. The ARL's are given in the following table, where $\delta = (\mu - \mu_0) / (\sigma_0/\sqrt{n})$.

δ	EWMA - ARL, $r = 0.08$, $h = 2.415$	DEWMA - ARL, $r_1 = 0.2$, $r_2 = 0.4$
0.0	199.10	198.11
0.1	139.13	154.42
0.2	74.19	91.13
0.3	42.57	53.49
0.4	27.24	33.63
0.5	19.11	22.97
0.6	14.21	16.49
0.7	11.00	12.58
0.8	8.89	9.93
0.9	7.35	8.11
1.0	6.18	6.78

Table 5.1: Comparison of EWMA and DEWMA $\overline{X}$ Charts

As one can see from the table the EWMA $\overline{X}$ chart with variable control limits having chart parameters $r = r_1 r_2 = 0.08$ and $h = 2.415$ outperforms the double EWMA $\overline{X}$ chart with variable control limits having chart parameters $r_1 = 0.2$, $r_2 = 0.4$, and $h = 2.44$ for values of $\delta \leq 1$.

Similarly, simulation was used to estimate the ARL's the double EWMA $\overline{X}$ chart with variable control limits having chart parameters $r_1 = 0.4$, $r_2 = 0.2$, and $h = 2.385$ for $\delta = 0.0, 0.1, ..., 1.0$. The in-control ARL of the chart was estimated to be 197.96. For both charts, fifty thousand $(50, 000)$ simulations were used.

The ARL's are given in the table that follows. The EWMA $\overline{X}$ chart with $r = r_1 r_2$ can be designed to have better run length performance than the DEWMA chart.

δ	EWMA - ARL, $r = 0.08$, $h = 2.415$	DEWMA - ARL, $r_1 = 0.4$, $r_2 = 0.2$
0.0	199.10	197.96
0.1	139.13	153.84
0.2	74.19	90.50
0.3	42.57	53.70
0.4	27.24	33.69
0.5	19.11	21.78
0.6	14.21	16.54
0.7	11.00	12.58
0.8	8.89	9.92
0.9	7.35	8.06
1.0	6.18	6.76

Table 5.2: Comparison of EWMA and DEWMA $\overline{X}$ Charts

Using Method 1, we further see from the following table that the EWMA $\overline{X}$ Chart performs better than the DEWMA $\overline{X}$ Chart.

δ	EWMA - ARL, $r = 0.04$, $h = 2.16$	DEWMA - ARL, $r_1 = 0.2$, $r_2 = 0.2$
0.0	199.36	199.04
0.1	148.76	192.20
0.2	82.94	170.80
0.3	48.98	134.03
0.4	32.46	92.87
0.5	23.70	59.95
0.6	18.56	39.07
0.7	15.15	27.65
0.8	12.86	20.85
0.9	11.15	16.56
1.0	9.90	13.79

Table 5.3: Comparison of EWMA and DEWMA $\overline{X}$ Charts

Using Method 2, using simulation we give a comparison of the EWMA $\overline{X}$ chart and the DEWMA $\overline{X}$ chart for shifts $\delta = 0.0, \ldots, 1.0$ in the following table. For all shifts listed the EWMA $\overline{X}$ chart outperforms the DEWMA $\overline{X}$ chart.

δ	EWMA - ARL, $r = 0.1065$, $h = 2.505$	DEWMA - ARL, $r_1 = 0.2$, $r_2 = 0.2$
0.0	199.50	199.04
0.1	155.04	192.20
0.2	89.44	170.80
0.3	52.64	134.03
0.4	34.44	92.87
0.5	24.59	59.95
0.6	18.86	39.07
0.7	15.24	27.65
0.8	12.72	20.85
0.9	10.94	16.56
1.0	9.62	13.79

Table 5.4: Comparison of EWMA and DEWMA $\overline{X}$ Charts

Rigdon (2023) via simulation has shown that the optimal double EWMA (EWMA $^{(2)}$) $\overline{X}$ chart is the optimal EWMA $\overline{X}$ chart. We conjecture that his method could be used to show that the optimal triple EWMA (EWMA $^{(3)}$) $\overline{X}$ chart is the optimal EWMA $\overline{X}$ chart. This would require a sizeable amount of simulations. Similarly, we conjecture that the optimal quadruple EWMA (EWMA $^{(4)}$) $\overline{X}$ chart is the optimal EWMA $\overline{X}$ chart.

5.7 Conclusion

We have shown that ad hoc double EWMA $\overline{X}$ chart does not perform as well as the EWMA $\overline{X}$ chart for small to moderate shifts in the process mean. Hence, the extra computations necessary to compute the plotted statistic as well as the extra work in setting up the control limits provide the practitioner no advantage. While not shown here, we do not believe that the triple, quadruple, etc. EWMA $\overline{X}$ charts will outperform the EWMA $\overline{X}$ chart for small to moderate shifts in the process mean. These charts would be even more difficult to manage by the practitioner than the double EWMA $\overline{X}$ chart.

CHAPTER 6

CHANGE POINT ANALYSIS

6.1 Introduction

When a production process has changed via one or more process parameters, it is desirable to know when this change has occurred or a least a prediction of when the change occurred. Our change point model assumes that the process changes from an in-control process after sampling stage τ but before sampling stage $\tau + 1$, where $0 \leq \tau < t$ with t the number of the sample in Phase II when the chart signals a potential out-of-control process. We will examine the likelihood method for the Shewhart and EWMA $\overline{X}$ charts to detect the change point of a production process. We assume the independent Normal model in Phase II. In the next section, we examine the likelihood function of associated with the Shewhart $\overline{X}$ both when the in-control parameters are known and when they are estimated. A method was derived for predicting the change point τ. In Section 3, we derive the likelihood function associated with the EWMA $\overline{X}$ chart in the parameters known case and the parameters estimated case. The method for predicting the change point τ is derived. This is followed by a section in which comparisons are made. In the last section, conclusions are given.

6.2 Predicting the Change Point Using a Shewhart $\overline{X}$ Chart

Walter A. Shewhart introduced the quality control chart in the early 1920's.The Shewhart $\overline{X}$ chart is a plot of the points $\left(t, \overline{X}_t\right)$ with control limits

$$LCL = \mu_0 - k\frac{\sigma_0}{\sqrt{n}} \text{ and } UCL = \mu_0 + k\frac{\sigma_0}{\sqrt{n}}.$$

The Shewhart $\overline{X}$ chart is a member of the family of EWMA $\overline{X}$ charts in which $r = 1$. The chart does not signal at time i if

$$\mu_0 - k\frac{\sigma_0}{\sqrt{n}} < \overline{X}_i < \mu_0 + k\frac{\sigma_0}{\sqrt{n}} \text{ or}$$
$$-k < \frac{\overline{X}_i - \mu_0}{\sigma_0/\sqrt{n}} < k \text{ or } -k < Y_t < k,$$

where

$$Y_i = \frac{\overline{X}_i - \mu_0}{\sigma_0/\sqrt{n}}.$$

The chart that plots the points with coordinates (i, Y_i) with control limits

$$LCL = -k \text{ and } UCL = k$$

is equivalent to the Shewhart $\overline{X}$ chart in the sense that either both charts signal or both charts do not signal at time t. The value of k for an in-control averge run length of ARL_0 must statisfy the equation

$$\frac{1}{2\left[1 - \Phi\left(k\right)\right]} = ARL_0 \text{ or } k = \Phi^{-1}\left(\frac{ARL_0 - 0.5}{ARL_0}\right).$$

For example, if we desire an in-control average run length of 200, we would select $k = 2.807033768$.

Assuming the process is in-control ($\mu = \mu_0$ and $\sigma = \sigma_0$) for $0 \le i \le \tau$ and the process is out-of-control with respect to the mean ($\mu \ne \mu_0$ and $\sigma = \sigma_0$) for $\tau < i \le t$, then we have that

$$Y_i \sim \begin{cases} N\left(0,1\right), & \text{for } 0 \le i \le \tau; \\ N\left(\delta,1\right), & \text{for } \tau < i \le t, \end{cases}$$

where

$$\delta = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}.$$

The likelihood function is the joint distribution of the Y_t's and is given by

$$L(\delta, \tau | \mathbf{Y}, k, t) = f_{\mathbf{Y}}(\mathbf{Y} | \delta, \tau, k, t) = \prod_{i=1}^{t} f_{Y_i}(Y_i | \delta, \tau, k, t),$$

where $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_t]^{\mathbf{T}}$. Under the independent Normal model with the process in-control for $0 \leq i \leq \tau$ and out-of-control for $\tau < i < t$, we have

$$\begin{aligned} L(\delta, \tau | \mathbf{Y}, k, t) &= f_{\mathbf{Y}}(\mathbf{Y} | \delta, \tau, k, t) \\ &= \prod_{i=1}^{\tau} \frac{1}{(2\pi)^{1/2}} e^{-\frac{1}{2}Y_i^2} \prod_{i=\tau+1}^{t} \frac{1}{(2\pi)^{1/2}} e^{-\frac{1}{2}(Y_i-\delta)^2} \\ &= \frac{1}{(2\pi)^{t/2}} e^{-\frac{1}{2}\sum_{i=1}^{\tau} Y_i^2} e^{-\frac{1}{2}\sum_{i=\tau+1}^{t}(Y_i-\delta)^2} \\ &= \frac{1}{(2\pi)^{t/2}} e^{-\frac{1}{2}\sum_{i=1}^{\tau} Y_i^2 - \frac{1}{2}\sum_{i=\tau+1}^{t}\left(Y_i^2 - 2\delta y_i + \delta^2\right)} \\ &= \frac{1}{(2\pi)^{t/2}} e^{-\frac{1}{2}\sum_{i=1}^{t} Y_i^2 + \delta\sum_{i=\tau+1}^{t} Y_i - \frac{1}{2}(t-\tau)\delta^2}. \end{aligned}$$

The log-likelihood function is given by

$$l(\delta, \tau | \mathbf{Y}, k, t) = -\frac{t}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{t} Y_i^2 + \delta\sum_{i=\tau+1}^{t} Y_i - \frac{1}{2}(t-\tau)\delta^2.$$

A method for predicting the change-point τ first uses the likelihood function to obtain the maximum likelihood estimator $\widehat{\delta}$ of δ. Taking the partial derivatives of $l(\delta, \tau | \mathbf{Y}, k, t)$ with respect to δ, we have

$$\frac{\partial l}{\partial \delta} = \sum_{i=\tau+1}^{t} y_i - (t-\tau)\delta.$$

The maximum likelihood estimator $\widehat{\delta}$ of δ is the solution to the equation

$$0 = \sum_{i=\tau+1}^{t} Y_i - (t-\tau)\widehat{\delta}.$$

It follows that

$$\widehat{\delta} = \frac{1}{t-\tau}\sum_{i=\tau+1}^{T} Y_i = \overline{Y}_{t-\tau}.$$

We see that

$$l\left(\widehat{\delta},\tau\,|\mathbf{Y},k,t\right) = -\frac{t}{2}\ln\left(2\pi\right)-\frac{1}{2}\sum\nolimits_{i=1}^{T}Y_i^2+\left(t-\tau\right)\overline{Y}_{T-\tau}^2-\frac{1}{2}\left(t-\tau\right)\overline{Y}_{t-\tau}^2$$
$$= -\frac{t}{2}\ln\left(2\pi\right)-\frac{1}{2}\sum\nolimits_{i=1}^{T}Y_i^2+\frac{1}{2}\left(t-\tau\right)\overline{Y}_{t-\tau}^2.$$

The maximum likelihood predictor $\widehat{\tau}$ of τ is the value of τ that maximizes

$$\left(t-\tau\right)\overline{Y}_{t-\tau}^2,$$

for $\tau = 0, 1, \ldots, t-1$.

In the following table, we have the change point predictions when in-control parameters are known using simulation based on $10,000$ simulated values for $\tau = 20, 40, 60$ and $\delta = 0.1, \ldots, 1.0$. Note that in all cases the predicted values of τ fell between τ and $\tau + 1$.

$\tau = 20$			$\tau = 40$			$\tau = 60$		
δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$
0.1	20.394	0.490	0.1	40.375	0.486	0.1	60.352	0.479
0.2	20.395	0.490	0.2	40.430	0.496	0.2	60.383	0.487
0.3	20.389	0.489	0.3	40.410	0.493	0.3	60.409	0.492
0.4	20.433	0.497	0.4	40.387	0.489	0.4	60.398	0.491
0.5	20.429	0.496	0.5	40.441	0.498	0.5	60.395	0.490
0.6	20.397	0.491	0.6	40.395	0.490	0.6	60.431	0.497
0.7	20.349	0.477	0.7	40.401	0.492	0.7	60.416	0.494
0.8	20.396	0.491	0.8	40.365	0.483	0.8	60.433	0.497
0.9	20.439	0.498	0.9	40.416	0.494	0.9	60.453	0.498
1.0	20.455	0.499	1.0	40.315	0.465	1.0	60.452	0.499

Table 6.1: In-control Known Paramters Case, ARL_0 equal to 200

The estimated parameter version of the chart plots the points with coordinates (i, Y_i^*) with control limits $LCL = -k$ and $UCL = k$, where

$$Y_i^* = c_{m,n} \frac{\overline{X}_i - \overline{\overline{X}}_0}{\overline{V}_0^{1/2} / \sqrt{n}}.$$

The value of $c_{m,n}$ is a function of m and n, where

$$c_{m,n} = \frac{\sqrt{2}\Gamma\left(\frac{m(n-1)}{2}\right)}{\sqrt{m\,(n-1)}\Gamma\left(\frac{m(n-1)-1}{2}\right)}.$$

The value of $c_{m,n}$ is selected so that $E\left(Y_i^* \left| \overline{X}_i\right.\right) = \left(\overline{X}_i - \mu_0\right) / (\sigma_0/\sqrt{n})$. Hence,

$$Y_i^* = c_{m,n} \frac{\overline{X}_i - \overline{\overline{X}}_0}{\overline{V}_0^{1/2} / \sqrt{n}} = \frac{\sqrt{2}\Gamma\left(\frac{m(n-1)}{2}\right)}{\sqrt{m\,(n-1)}\Gamma\left(\frac{m(n-1)-1}{2}\right)} \frac{\overline{X}_i - \overline{\overline{X}}_0}{\overline{V}_0^{1/2} / \sqrt{n}}.$$

Note that we can write Y_i^* as

$$Y_i^* = \frac{c_{m,n}}{\overline{V}_0^{1/2}/\sigma_0}\left(\frac{\overline{X}_i-\mu}{\sigma_0/\sqrt{n}}+\frac{\mu-\mu_0}{\sigma_0/\sqrt{n}}-\frac{\overline{\overline{X}}_0-\mu_0}{\sigma_0/\sqrt{mn}}/\sqrt{m}\right)$$
$$= \frac{c_{m,n}}{\overline{W}_0^{1/2}}\left(Z_i+\delta-\overline{Z}_0/\sqrt{m}\right) = c_{m,n}\overline{W}_0^{-1/2}\left(Z_i+\delta-\overline{Z}_0/\sqrt{m}\right),$$

where

$$Z_i = \frac{\overline{X}_i-\mu}{\sigma_0/\sqrt{n}} \sim N(0,1), \overline{W}_0 = \overline{V}_0/\sigma_0^2 \sim \frac{1}{m(n-1)}\chi^2_{m(n-1)},$$
$$\delta = \frac{\mu-\mu_0}{\sigma_0/\sqrt{n}}, \text{ and } \overline{Z}_0 = \frac{\overline{\overline{X}}_0-\mu_0}{\sigma_0/\sqrt{mn}} \sim N(0,1).$$

The conditional distribution of Y_i^* given the statistics $\overline{Z}_0$ and $\overline{W}_0$ is

$$Y_i^* \left|\overline{Z}_0, \overline{W}_0\right. \sim \begin{cases} N\left(-c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}, c^2\overline{W}_0^{-1}\right), & 0 \le i \le \tau; \\ N\left(\delta - c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}, c^2\overline{W}_0^{-1}\right), & \tau < i \le t, \end{cases}$$

The conditional likelihood function is

$$L\left(\delta,\tau\left|\mathbf{Y}^*,\overline{Z}_0,\overline{W}_0,k,t\right.\right) = \prod_{i=1}^{\tau}\frac{1}{(2\pi)^{1/2}c_{m,n}\overline{W}_0^{-1/2}}e^{-\frac{1}{2}\left(\frac{Y_i^*+c\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}}{c\overline{W}_0^{-1/2}}\right)^2}$$
$$\times\prod_{i=\tau+1}^{t}\frac{1}{(2\pi)^{1/2}c_{m,n}\overline{W}_0^{-1/2}}e^{-\frac{1}{2}\left(\frac{Y_i^*-\delta+c\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}}{c\overline{W}_0^{-1/2}}\right)^2}$$
$$= \frac{1}{(2\pi)^{t/2}\overline{W}_0^{-t/2}c_{m,n}^t}e^{-\frac{1}{2}c_{m,n}^{-2}\overline{W}_0\sum_{i=1}^{\tau}\left(Y_i^*+c\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)^2}$$
$$\times e^{-\frac{1}{2}c_{m,n}^{-2}\overline{W}_0\sum_{i=\tau+1}^{t}\left(Y_i^*+c\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}-\delta\right)^2},$$

where $\mathbf{Y}^* = [Y_1^*,\ldots,Y_t^*]^{\mathbf{T}}$. Observe that

$$\left(Y_i^*-\delta+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)^2 = \left(Y_i^*+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}-\delta\right)^2$$
$$= \left(Y_i^*+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)^2 - 2\left(Y_i^*+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)\delta+\delta^2.$$

Thus,

$$\begin{aligned}
&L\left(\delta,\tau\left|\mathbf{Y}^{*},\overline{Z}_0,\overline{W}_0,k,t\right.\right)\\
&=\frac{1}{(2\pi)^{t/2}\,c_{m,n}^{t}\overline{W}_0^{-t/2}}e^{-\frac{1}{2}c_{m,n}^{-2}\overline{W}_0\sum_{i=1}^{t}\left(Y_i^{*}+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)^2}\\
&\times e^{c_{m,n}^{-2}\overline{W}_0\sum_{i=\tau+1}^{t}\left(Y_i^{*}+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)\delta-\frac{1}{2}c_{m,n}^{-2}\overline{W}_0\sum_{i=\tau+1}^{T}(t-\tau)\delta^2}.
\end{aligned}$$

We can now write the conditional log-likelihood function as

$$\begin{aligned}
l\left(\delta,\tau\left|\mathbf{Y}^{*},\overline{Z}_0,\overline{W}_0,k,t\right.\right)&=-\frac{t}{2}\ln(2\pi)-t\ln(c_{m,n})+\frac{t}{2}\ln\left(\overline{W}_0\right)\\
&-\frac{1}{2}c_{m,n}^{-2}\overline{W}_0\sum\nolimits_{i=1}^{t}\left(Y_i^{*}+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)^2\\
&+c_{m,n}^{-2}\overline{W}_0\sum\nolimits_{i=\tau+1}^{t}\left(Y_i^{*}+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)\delta\\
&-\frac{1}{2}c_{m,n}^{-2}\overline{W}_0\,(t-\tau)\,\delta^2.
\end{aligned}$$

Observe that

$$\frac{\partial l}{\partial\delta}=c_{m,n}^{-2}\overline{W}_0\sum\nolimits_{i=\tau+1}^{t}\left(Y_i^{*}+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)-(t-\tau)\,c_{m,n}^{-2}\overline{W}_0\delta.$$

The maximum likelihood estimator $\widehat{\delta}$ of δ is the solution to the equation

$$c_{m,n}^{-2}\overline{W}_0\sum\nolimits_{i=\tau+1}^{t}\left(Y_i^{*}+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}\right)-(t-\tau)\,c_{m,n}^{-2}\overline{W}_0\delta=0.\text{or }\widehat{\delta}=\overline{Y}_{t-\tau}^{*}+c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}/\sqrt{m}.$$

The conditional log-likelihood function evaluated at $\widehat{\delta}$ yields

$$\begin{aligned}
l\left(\widehat{\delta},\tau\left|\mathbf{Y}^{*},\overline{Z}_{0},\overline{W}_{0},k,t\right.\right) &= -\frac{t}{2}\ln\left(2\pi\right)-t\ln\left(c_{m,n}\right) \\
&+\frac{t}{2}\ln\left(\overline{W}_{0}\right)-\frac{1}{2}c_{m,n}^{-2}\overline{W}_{0}\sum\nolimits_{i=1}^{T}\left(Y_{i}^{*}+c_{m,n}\overline{Z}_{0}\overline{W}_{0}^{-1/2}/\sqrt{m}\right)^{2} \\
&+c^{-2}\overline{W}_{0}\left(t-\tau\right)\left(\overline{Y}_{t-\tau}^{*}+c_{m,n}\overline{Z}_{0}\overline{W}_{0}^{-1/2}/\sqrt{m}\right)^{2} \\
&-\frac{1}{2}c_{m,n}^{-2}\overline{W}_{0}\left(t-\tau\right)\left(\overline{Y}_{t-\tau}^{*}+c_{m,n}\overline{Z}_{0}\overline{W}_{0}^{-1/2}/\sqrt{m}\right)^{2} \\
&= -\frac{t}{2}\ln\left(2\pi\right)-t\ln\left(c_{m,n}\right) \\
&+\frac{t}{2}\ln\left(\overline{W}_{0}\right)-\frac{1}{2}c_{m,n}^{-2}\overline{W}_{0}\sum\nolimits_{i=1}^{t}\left(Y_{i}^{*}+c_{m,n}\overline{Z}_{0}\overline{W}_{0}^{-1/2}/\sqrt{m}\right)^{2} \\
&+\frac{1}{2}c_{m,n}^{-2}\overline{W}_{0}\left(t-\tau\right)\left(\overline{Y}_{t-\tau}^{*}+c_{m,n}\overline{Z}_{0}\overline{W}_{0}^{-1/2}/\sqrt{m}\right)^{2}.
\end{aligned}$$

The maximum likelihood predictor $\widehat{\tau}$ of τ is the value of τ that maximizes

$$\left(t-\tau\right)\left(\overline{Y}_{t-\tau}^{*}+c_{m,n}\overline{Z}_{0}\overline{W}_{0}^{-1/2}/\sqrt{m}\right)^{2}.$$

Note that if $c_{m,n}=1$, $\overline{Z}_{0}=0$, and $\overline{W}_{0}=1$, yields the method for obtaining the maximum likelihood predictor $\widehat{\tau}$ of τ when the in-control parameters are given.

In the following table, we have the change point predictions when in-control parameters are estimated using simulation based on $10,000$ simulated values for $\tau=20,40,60$ and $\delta=0.1,\ldots,1.0$. The values for m and n were set to 4 and 6, respectively. Note that in all cases the predicted values of τ fell between τ and $\tau+1$. Also in a comparison with the known parameters case, the predicted values are a bit larger with more variability.

$\tau = 20$			$\tau = 40$			$\tau = 60$		
δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$
0.1	20.520	0.500	0.1	40.515	0.500	0.1	60.459	0.498
0.2	20.515	0.501	0.2	40.501	0.501	0.2	60.505	0.501
0.3	20.493	0.500	0.3	40.526	0.499	0.3	60.489	0.500
0.4	20.580	0.494	0.4	40.563	0.497	0.4	60.541	0.499
0.5	20.500	0.501	0.5	40.547	0.498	0.5	60.551	0.498
0.6	20.585	0.493	0.6	40.577	0.495	0.6	60.607	0.489
0.7	20.616	0.487	0.7	40.659	0.475	0.7	60.580	0.494
0.8	20.667	0.472	0.8	40.607	0.489	0.8	60.664	0.473
0.9	20.698	0.460	0.9	40.683	0.466	0.9	60.687	0.464
1.0	20.706	0.457	1.0	40.664	0.473	1.0	60.662	0.474

Table 6.2: In-control Estimated Parameters Case, ARL_0 equal to 200

6.3 Change Point Prediction Using a EWMA $\overline{X}$ Chart

In the parameters known case, the $t \times 1$ vector $\mathbf{U}$ of EWMA statistics can in general be expressed as

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_{\tau} \\ \mathbf{U}_{t-\tau} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{\tau} & \mathbf{0} \\ \mathbf{J} & \mathbf{L}_{t-\tau} \end{bmatrix} \begin{bmatrix} \mathbf{Z}_{\tau} \\ \mathbf{Z}_{t-\tau} \end{bmatrix} + \delta \begin{bmatrix} \mathbf{0}_{\tau} \\ \mathbf{b}_{t-\tau} \end{bmatrix}$$
$$\sim N_t \left(\mu_{\mathbf{U}} = \delta \begin{bmatrix} \mathbf{0}_{\tau} \\ \mathbf{b}_{t-\tau} \end{bmatrix}, \mathbf{\Sigma}_{\mathbf{U}} = \begin{bmatrix} \mathbf{L}_{\tau} & \mathbf{0} \\ \mathbf{J} & \mathbf{L}_{t-\tau} \end{bmatrix} \begin{bmatrix} \mathbf{L}_{\tau} & \mathbf{0} \\ \mathbf{J} & \mathbf{L}_{t-\tau} \end{bmatrix}^{\mathbf{T}} \right)$$

where $\mathbf{U}_\tau = [U_1, \ldots, U_\tau]^{\mathbf{T}}$, $\mathbf{U}_{t-\tau} = [U_{\tau+1}, \ldots, U_t]^{\mathbf{T}}$,

$$\mathbf{L}_\tau = \begin{bmatrix} r & 0 & \ldots & 0 \\ (1-r)\,r & r & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (1-r)^{\tau-1}\,r & (1-r)^{\tau-2}\,r & \ldots & r \end{bmatrix}^{\tau\times\tau},$$

$$\mathbf{J} = \begin{bmatrix} (1-r)^{\tau}\,r & (1-r)^{\tau-1}\,r & \ldots & (1-r)^{1}\,r \\ (1-r)^{\tau+1}\,r & (1-r)^{\tau}\,r & \ldots & (1-r)^{2}\,r \\ \vdots & \vdots & \ddots & \vdots \\ (1-r)^{\tau+t-\tau-1}\,r & (1-r)^{\tau+t-\tau-2}\,r & \ldots & (1-r)^{t-\tau}\,r \end{bmatrix}^{(t-\tau)\times\tau},$$

$$\mathbf{L}_{t-\tau} = \begin{bmatrix} r & 0 & \ldots & 0 \\ (1-r)\,r & r & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (1-r)^{t-1}\,r & (1-r)^{t-2}\,r & \ldots & r \end{bmatrix}^{(t-\tau)\times(t-\tau)},$$

and $$\mathbf{b}_{t-\tau} = \begin{bmatrix} 1-(1-r)^{1} \\ 1-(1-r)^{2} \\ \vdots \\ 1-(1-r)^{t-\tau} \end{bmatrix}.$$

The likelihood function $L\left(\delta, \tau \,|\mathbf{U}, r, h, t\right)$ as a function of δ and τ is

$$\begin{aligned} L\left(\delta, \tau \,|\mathbf{U}, r, h, t\right) &= f_{\mathbf{U}}\left(\mathbf{U}\,|\delta, \tau, r, h, t\right) = \frac{1}{(2\pi)^{t/2}\left|\boldsymbol{\Sigma}_{\mathbf{U}}\right|^{1/2}} e^{-\frac{1}{2}(\mathbf{U}-\mu_{\mathbf{U}})^{\mathbf{T}}\boldsymbol{\Sigma}_{\mathbf{U}}^{-1}(\mathbf{U}-\mu_{\mathbf{U}})} \\ &= \frac{1}{(2\pi)^{t/2}\,r^t} e^{-\frac{1}{2}(\mathbf{U}-\mu_{\mathbf{U}})^{\mathbf{T}}\boldsymbol{\Sigma}_{\mathbf{U}}^{-1}(\mathbf{U}-\mu_{\mathbf{U}})} = \frac{1}{(2\pi)^{t/2}\,r^t} e^{-\frac{1}{2}\mathbf{Z}^{\mathbf{T}}\mathbf{Z}} \text{ and} \\ l\left(\delta, \tau \,|\mathbf{U}, r, h, t\right) &= -\frac{t}{2}\ln(2\pi) - t\ln(r) - \frac{1}{2}\mathbf{Z}^{\mathbf{T}}\mathbf{Z}, \end{aligned}$$

where

$$\left(\mathbf{U}-\mu_{\mathbf{U}}\right)^{\mathbf{T}} \mathbf{\Sigma}_{\mathbf{U}}^{-1}\left(\mathbf{U}-\mu_{\mathbf{U}}\right)=\mathbf{Z}^{\mathbf{T}}\mathbf{Z},$$

with

$$\begin{aligned}
\mathbf{Z} &= \begin{bmatrix} \mathbf{Z}_{\tau} \\ \mathbf{Z}_{t-\tau} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{L}_{\tau} & \mathbf{0} \\ \mathbf{J} & \mathbf{L}_{t-\tau} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{U}_{\tau} \\ \mathbf{U}_{t-\tau} \end{bmatrix} - \delta \begin{bmatrix} \mathbf{0}_{\tau} \\ \mathbf{b}_{t-\tau} \end{bmatrix} \right) \\
&= \begin{bmatrix} \mathbf{L}_{\tau}^{-1} & \mathbf{0} \\ \mathbf{K} & \mathbf{L}_{t-\tau}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\tau} \\ \mathbf{U}_{t-\tau} \end{bmatrix} - \delta \begin{bmatrix} \mathbf{L}_{\tau}^{-1} & \mathbf{0} \\ \mathbf{K} & \mathbf{L}_{t-\tau}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0}_{\tau} \\ \mathbf{b}_{t-\tau} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{L}_{\tau}^{-1} & \mathbf{0} \\ \mathbf{K} & \mathbf{L}_{t-\tau}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\tau} \\ \mathbf{U}_{t-\tau} \end{bmatrix} - \delta \begin{bmatrix} \mathbf{0}_{\tau} \\ \mathbf{1}_{t-\tau} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{L}_{\tau}^{-1}\mathbf{U}_{\tau} \\ \mathbf{K}\mathbf{U}_{\tau} + \mathbf{L}_{t-\tau}^{-1}\mathbf{U}_{t-\tau} - \delta \mathbf{1}_{t-\tau} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \\ \mathbf{H} - \delta \mathbf{1}_{t-\tau} \end{bmatrix}
\end{aligned}$$

$\mathbf{G} = \mathbf{L}_{\tau}^{-1}\mathbf{U}_{\tau}$, $\mathbf{H} = \mathbf{K}\mathbf{U}_{\tau} + \mathbf{L}_{t-\tau}^{-1}\mathbf{U}_{t-\tau}$ and

$$\mathbf{K} = \begin{bmatrix} 0 & \dots & 0 & -\frac{1-r}{r} \\ 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix}.$$

It follows that

$$
\begin{aligned}
-\frac{1}{2}\mathbf{Z}^{\mathbf{T}}\mathbf{Z} &= -\frac{1}{2}\begin{bmatrix} \mathbf{G} \\ \mathbf{H}-\delta\mathbf{1}_{t-\tau} \end{bmatrix}^{\mathbf{T}} \begin{bmatrix} \mathbf{G} \\ \mathbf{H}-\delta\mathbf{1}_{t-\tau} \end{bmatrix} \\
&= -\frac{1}{2}\mathbf{G}^{\mathbf{T}}\mathbf{G} - \frac{1}{2}\left(\mathbf{H}-\delta\mathbf{1}_{t-\tau}\right)^{\mathbf{T}}\left(\mathbf{H}-\delta\mathbf{1}_{t-\tau}\right) \\
&= -\frac{1}{2}\left(\mathbf{G}^{\mathbf{T}}\mathbf{G}+\mathbf{H}^{\mathbf{T}}\mathbf{H}\right) + \mathbf{H}^{\mathbf{T}}\mathbf{1}_{t-\tau}\delta - \frac{1}{2}\mathbf{1}_{t-\tau}^{\mathbf{T}}\mathbf{1}_{t-\tau}\delta^2 \\
&= -\frac{1}{2}\left(\mathbf{G}^{\mathbf{T}}\mathbf{G}+\mathbf{H}^{\mathbf{T}}\mathbf{H}\right) + \mathbf{H}^{\mathbf{T}}\mathbf{1}_{t-\tau}\delta - \frac{1}{2}\left(t-\tau\right)\delta^2.
\end{aligned}
$$

Hence, we can write the log-likelihood function as

$$
l\left(\delta,\tau\,|\mathbf{U},r,h,t\right) = -\frac{t}{2}\ln\left(2\pi\right) - t\ln\left(r\right) - \frac{1}{2}\left(\mathbf{G}^{\mathbf{T}}\mathbf{G}+\mathbf{H}^{\mathbf{T}}\mathbf{H}\right) + \mathbf{H}^{\mathbf{T}}\mathbf{1}_{t-\tau}\delta - \frac{1}{2}\left(t-\tau\right)\delta^2.
$$

The maximum likelihood estimator $\widehat{\delta}$ of δ is the solution to the equation

$$
\frac{\partial l}{\partial \delta} = \mathbf{H}^{\mathbf{T}}\mathbf{1}_{t-\tau} - \left(t-\tau\right)\delta = 0 \text{ or } \widehat{\delta} = \frac{\mathbf{H}^{\mathbf{T}}\mathbf{1}_{t-\tau}}{t-\tau}.
$$

Recall that $\mathbf{H} = \mathbf{K}\mathbf{U}_\tau + \mathbf{L}_{t-\tau}^{-1}\mathbf{U}_{t-\tau}$, thus

$$
\begin{aligned}
\mathbf{H}^{\mathbf{T}}\mathbf{1}_{t-\tau} &= \left(\mathbf{K}\mathbf{U}_\tau + \mathbf{L}_{t-\tau}^{-1}\mathbf{U}_{t-\tau}\right)^{\mathbf{T}}\mathbf{1}_{t-\tau} \\
&= \mathbf{U}_\tau^{\mathbf{T}}\mathbf{K}^{\mathbf{T}}\mathbf{1}_{t-\tau} + \mathbf{U}_{t-\tau}^{\mathbf{T}}\left(\mathbf{L}_{t-\tau}^{-1}\right)^{\mathbf{T}}\mathbf{1}_{t-\tau} \\
&= \left(t-\tau\right)\left[\frac{1}{t-\tau}\sum\nolimits_{i=\tau}^{t-1}U_i + \frac{1}{r}\left(\frac{U_t-U_\tau}{t-\tau}\right)\right] \\
&= \left(t-\tau\right)\left[\frac{1}{t-\tau}\sum\nolimits_{i=\tau}^{t-1}U_i + \frac{1}{r}\left(\frac{U_t-U_\tau}{t-\tau}\right)\right].
\end{aligned}
$$

It follows that the likelihood function can be expressed as

$$
L\left(\delta,\tau\,|\mathbf{U},r,h,t\right) = \frac{1}{\left(2\pi\right)^{t/2}r^{t/2}}e^{A+B\delta-\frac{1}{2}(t-\tau)\delta^2}.
$$

Hence, the log-likelihood function can be expressed as

$$
l\left(\delta,\tau\,|\mathbf{U},r,h,t\right) = -\frac{t}{2}\ln\left(2\pi\right) - \frac{t}{2}\ln\left(r\right) + A + B\delta - \frac{1}{2}\left(t-\tau\right)\delta^2.
$$

The maximum likelihood estimator $\widehat{\delta}$ of δ is the solution to the equation

$$\begin{aligned}\frac{\partial l}{\partial \delta} &= B - (t-\tau)\,\delta = 0 \text{ or} \\ \widehat{\delta} &= \frac{1}{t-\tau}B = \frac{1}{t-\tau}\left(\mathbf{1}^{\mathbf{T}}_{t-\tau}\mathbf{K}\mathbf{U}_{\tau} + \mathbf{1}^{\mathbf{T}}_{t-\tau}\mathbf{L}^{-1}_{t-\tau}\mathbf{U}_{t-\tau}\right) \\ &= \frac{1}{t-\tau}\left(-\frac{1-r}{r}U_{\tau} + \sum\nolimits_{i=\tau+1}^{t} U_i + \frac{1-r}{r}U_t\right) \\ &= \frac{1}{t-\tau}\left[\sum\nolimits_{i=\tau+1}^{t} U_i + \frac{1-r}{r}\left(U_t - U_{\tau}\right)\right] \\ &= \overline{U}_{t-\tau} + \frac{1-r}{r}\left(\frac{U_t - U_{\tau}}{t-\tau}\right).\end{aligned}$$

Evaluating the log-likelihood function at $\widehat{\delta}$, we have

$$l\left(\widehat{\delta},\tau\,|\mathbf{U},r,h,t\right) = -\frac{t}{2}\ln(2\pi) - \frac{t}{2}\ln(r) + A + \frac{1}{2}(t-\tau)\left[\overline{U}_{t-\tau} + \frac{1-r}{r}\left(\frac{U_t - U_{\tau}}{t-\tau}\right)\right]^2.$$

$$\begin{aligned}l\left(\widehat{\delta},\tau\,|\mathbf{U},r,h,t\right) = &-\frac{t}{2}\ln(2\pi) - \frac{t}{2}\ln(r) + A \\ &+ \frac{1}{2}(t-\tau)\left[\overline{U}_{t-\tau} + \frac{1-r}{r}\left(\frac{U_t - U_{\tau}}{t-\tau}\right)\right]^2.\end{aligned}$$

The maximum likelihood predictor $\widehat{\tau}$ for τ is the value of τ that maximizes

$$(t-\tau)\left(\overline{U}_{t-\tau} + \frac{1-r}{r}\frac{U_t - U_{\tau}}{t-\tau}\right)^2$$

for $\tau = 0, 1, \ldots, t-1$.

In the following table, we have the change point predictions when in-control parameters are estimated using simulation based on $10,000$ simulated values for $\tau = 20, 40, 60$ and $\delta = 0.1, \ldots, 1.0$. The values for m and n were set to 4 and 6, respectively. Note that in all cases the predicted values of τ fell between τ and $\tau + 1$. Also in a comparison with the Shewhart $\overline{X}$ chart with known in-control parameters, the predicted values are a bit larger with more variability.

$\tau = 20$			$\tau = 40$			$\tau = 60$		
δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$
0.1	20.607	0.489	0.1	40.617	0.487	0.1	60.576	0.485
0.2	20.603	0.489	0.2	40.618	0.486	0.2	60.630	0.483
0.3	20.657	0.475	0.3	40.643	0.479	0.3	60.637	0.481
0.4	20.671	0.470	0.4	40.672	0.469	0.4	60.681	0.466
0.5	20.703	0.457	0.5	40.698	0.459	0.5	60.704	0.457
0.6	20.738	0.440	0.6	40.730	0.444	0.6	60.732	0.443
0.7	20.757	0.429	0.7	40.776	0.417	0.7	60.762	0.426
0.8	20.798	0.402	0.8	40.810	0.393	0.8	60.806	0.395
0.9	20.832	0.374	0.9	40.845	0.362	0.9	60.837	0.370
1.0	20.860	0.347	1.0	40.852	0.355	1.0	60.864	0.342

Table 6.3: In-control Known Parameters Case, ARL_0 equal to 200

In the estimated parameters case, the $t \times 1$ vector $\mathbf{U}$ of EWMA statistics can be expressed as

$$\mathbf{U}^* = \begin{bmatrix} \mathbf{U}^*_\tau \\ \mathbf{U}^*_{t-\tau} \end{bmatrix} = c^2_{m,n}\overline{W}_0^{-1} \begin{bmatrix} \mathbf{L}_\tau & \mathbf{0} \\ \mathbf{J} & \mathbf{L}_{t-\tau} \end{bmatrix} \begin{bmatrix} \mathbf{Z}_\tau \\ \mathbf{Z}_{t-\tau} \end{bmatrix} - \frac{c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}}{\sqrt{m}} \begin{bmatrix} \mathbf{a}_\tau \\ \mathbf{a}_{t-\tau} \end{bmatrix} + c_{m,n}\overline{Z}_0\overline{W}_0^{-1/2}\delta \begin{bmatrix} \mathbf{0}_\tau \\ \mathbf{b}_{t-\tau} \end{bmatrix},$$

where

$$\mathbf{a}_\tau^{\tau\times 1} = [1, 2, \ldots, \tau]^{\mathbf{T}}, \ \mathbf{a}_{t-\tau}^{(t-\tau)\times 1} = [\tau+1, \tau+2, \ldots, t]^{\mathbf{T}}, \text{ and}$$

$$\mathbf{b}_{t-\tau}^{(t-\tau)\times 1} = [1, 2, \ldots, t-\tau]^{\mathbf{T}}.$$

Here we are defining $\mathbf{U}^*$ in terms of

$$Y_i^* = c_{m,n}\frac{\overline{X}_i - \overline{\overline{X}}_0}{\overline{V}_0^{1/2}/\sqrt{n}} = c_{m,n}\overline{W}_0^{-1/2}\left(Z_i + \delta - \overline{Z}_0/\sqrt{m}\right).$$

Under our model, it follows that

$$\mathbf{U}^* = \begin{bmatrix} \mathbf{U}_\tau^* \\ \mathbf{U}_{t-\tau}^* \end{bmatrix} \sim N_t\left(\mu_{\mathbf{U}}, \mathbf{\Sigma}_{\mathbf{U}}\right),$$

where

$$\mu_{\mathbf{U}} = -\frac{\overline{Z}_0\overline{W}_0^{-1/2}}{\sqrt{m}}\begin{bmatrix} \mathbf{a}_\tau \\ \mathbf{a}_{t-\tau} \end{bmatrix} + \delta\begin{bmatrix} \mathbf{0}_\tau \\ \mathbf{b}_{t-\tau} \end{bmatrix} \text{ and}$$

$$\mathbf{\Sigma}_{\mathbf{U}} = \overline{W}_0^{-1}\begin{bmatrix} \mathbf{L}_\tau & \mathbf{0} \\ \mathbf{J} & \mathbf{L}_{t-\tau} \end{bmatrix}\begin{bmatrix} \mathbf{L}_\tau & \mathbf{0} \\ \mathbf{J} & \mathbf{L}_{t-\tau} \end{bmatrix}^{\mathbf{T}}.$$

The likelihood function associated with the distribution of $\mathbf{U}$ is

$$L\left(\delta,\tau\,\middle|\,\mathbf{U}^*,\overline{Z}_0,\overline{W}_0,r,h,t\right) = \frac{1}{(2\pi)^{t/2}\,r^{t/2}}e^{-\frac{1}{2}(\mathbf{U}-\mu_{\mathbf{U}})^{\mathbf{T}}\mathbf{\Sigma}_{\mathbf{U}}^{-1}(\mathbf{U}-\mu_{\mathbf{U}})}$$

with log-likelihood function

$$\begin{aligned} l\left(\delta,\tau\,\middle|\,\mathbf{U}^*,\overline{Z}_0,\overline{W}_0,r,h,t\right) &= -\frac{t}{2}\ln(2\pi) - \frac{t}{2}\ln(r) - \frac{1}{2}\left(\mathbf{U}-\mu_{\mathbf{U}}\right)^{\mathbf{T}}\mathbf{\Sigma}_{\mathbf{U}}^{-1}\left(\mathbf{U}-\mu_{\mathbf{U}}\right) \\ &= -\frac{t}{2}\ln(2\pi) - \frac{t}{2}\ln(r) - \frac{1}{2}\mathbf{Z}^{\mathbf{T}}\mathbf{Z}, \end{aligned}$$

where

$$\begin{aligned} \begin{bmatrix} \mathbf{Z}_\tau \\ \mathbf{Z}_{t-\tau} \end{bmatrix} &= \begin{bmatrix} \overline{W}_0^{1/2}\mathbf{L}_\tau^{-1}\mathbf{U}_\tau + \frac{\overline{Z}_0}{\sqrt{m}}\mathbf{1}_\tau \\ \overline{W}_0^{1/2}\mathbf{K}\mathbf{U}_\tau + \overline{W}_0^{1/2}\mathbf{L}_{t-\tau}^{-1}\mathbf{U}_{t-\tau} + \frac{\overline{Z}_0}{\sqrt{m}}\mathbf{1}_{t-\tau} - \delta\mathbf{1}_{t-\tau} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{G} \\ \mathbf{H} - \delta\mathbf{1}_{t-\tau} \end{bmatrix}, \end{aligned}$$

with

$$\mathbf{G} = \overline{W}_0^{1/2}\mathbf{L}_\tau^{-1}\mathbf{U}_\tau + \frac{\overline{Z}_0}{\sqrt{m}}\mathbf{1}_\tau \text{ and } \mathbf{H} = \overline{W}_0^{1/2}\mathbf{K}\mathbf{U}_\tau + \overline{W}_0^{1/2}\mathbf{L}_{t-\tau}^{-1}\mathbf{U}_{t-\tau} + \frac{\overline{Z}_0}{\sqrt{m}}\mathbf{1}_{t-\tau}.$$

It follows that

$$\begin{aligned}
-\frac{1}{2}\mathbf{Z}^\mathbf{T}\mathbf{Z} &= -\frac{1}{2}\begin{bmatrix} \mathbf{G} \\ \mathbf{H} - \delta\mathbf{1}_{t-\tau} \end{bmatrix}^\mathbf{T}\begin{bmatrix} \mathbf{G} \\ \mathbf{H} - \delta\mathbf{1}_{t-\tau} \end{bmatrix} \\
&= -\frac{1}{2}\mathbf{G}^\mathbf{T}\mathbf{G} - \frac{1}{2}\left(\mathbf{H} - \delta\mathbf{1}_{t-\tau}\right)^\mathbf{T}\left(\mathbf{H} - \delta\mathbf{1}_{t-\tau}\right) \\
&= -\frac{1}{2}\mathbf{G}^\mathbf{T}\mathbf{G} - \frac{1}{2}\left(\mathbf{H}^\mathbf{T} - \delta\mathbf{1}_{t-\tau}^\mathbf{T}\right)\left(\mathbf{H} - \delta\mathbf{1}_{t-\tau}\right) \\
&= -\frac{1}{2}\mathbf{G}^\mathbf{T}\mathbf{G} - \frac{1}{2}\left(\mathbf{H}^\mathbf{T}\mathbf{H} - 2\delta\mathbf{H}^\mathbf{T}\mathbf{1}_{t-\tau}^\mathbf{T}\mathbf{1}_{t-\tau} + \delta^2\mathbf{1}_{t-\tau}^\mathbf{T}\mathbf{1}_{t-\tau}\right) \\
&= -\frac{1}{2}\mathbf{G}^\mathbf{T}\mathbf{G} - \frac{1}{2}\mathbf{H}^\mathbf{T}\mathbf{H} + \delta\mathbf{H}^\mathbf{T}\mathbf{1}_{t-\tau}^\mathbf{T}\mathbf{1}_{t-\tau} - \frac{1}{2}\delta^2\mathbf{1}_{t-\tau}^\mathbf{T}\mathbf{1}_{t-\tau}.
\end{aligned}$$

Hence, the log-likelihood function can be expressed as

$$\begin{aligned}
l\left(\delta, \tau \,\middle|\, \mathbf{U}^*, \overline{Z}_0, \overline{W}_0, r, h, t\right) &= -\frac{t}{2}\ln\left(2\pi\right) - \frac{t}{2}\ln\left(r\right) \\
&\quad - \frac{1}{2}\mathbf{G}^\mathbf{T}\mathbf{G} - \frac{1}{2}\mathbf{H}^\mathbf{T}\mathbf{H} + \delta\mathbf{H}^\mathbf{T}\mathbf{1}_{t-\tau} - \frac{1}{2}\delta^2\mathbf{1}_{t-\tau}^\mathbf{T}\mathbf{1}_{t-\tau}.
\end{aligned}$$

It follows that

$$\frac{\partial l}{\partial \delta} = \mathbf{H}^\mathbf{T}\mathbf{1}_{t-\tau} - \delta\mathbf{1}_{t-\tau}^\mathbf{T}\mathbf{1}_{t-\tau} = \mathbf{H}^\mathbf{T}\mathbf{1}_{t-\tau} - \delta\left(t-\tau\right).$$

The maximum likelihood estimator $\widehat{\delta}$ of δ is the solution to the equation

$$\mathbf{H}^\mathbf{T}\mathbf{1}_{t-\tau} - \widehat{\delta}\left(t-\tau\right) = 0 \text{ or } \widehat{\delta} = \frac{1}{t-\tau}\mathbf{H}^\mathbf{T}\mathbf{1}_{t-\tau}.$$

We see that

$$\begin{aligned}
\mathbf{H}^{\mathbf{T}}\mathbf{1}_{t-\tau} &= \overline{W}_0^{1/2}\left(\mathbf{K}\mathbf{U}_\tau + \mathbf{L}_{t-\tau}^{-1}\mathbf{U}_{t-\tau}\right)^{\mathbf{T}}\mathbf{1}_{t-\tau} \\
&= \overline{W}_0^{1/2}\left(\mathbf{U}_\tau^{\mathbf{T}}\mathbf{K}^{\mathbf{T}} + \mathbf{U}_{t-\tau}^{\mathbf{T}}\left(\mathbf{L}_{t-\tau}^{-1}\right)^{\mathbf{T}}\right)\mathbf{1}_{t-\tau} \\
&= \overline{W}_0^{1/2}\left(\mathbf{U}_\tau^{\mathbf{T}}\mathbf{K}^{\mathbf{T}}\mathbf{1}_{t-\tau} + \mathbf{U}_{t-\tau}^{\mathbf{T}}\left(\mathbf{L}_{t-\tau}^{-1}\right)^{\mathbf{T}}\mathbf{1}_{t-\tau}\right) \\
&= \overline{W}_0^{1/2}\left(-\frac{1-r}{r}U_\tau + \frac{1}{r}\sum\nolimits_{i=1}^{t} U_i - \frac{1-r}{r}\sum\nolimits_{i=2}^{t} U_i\right) \\
&= \overline{W}_0^{1/2}\left(-\frac{1-r}{r}U_\tau + \frac{1}{r}\sum\nolimits_{i=\tau+1}^{t} U_i - \frac{1-r}{r}\sum\nolimits_{i=\tau+1}^{t} U_i + \frac{1-r}{r}U_{\tau+1}\right) \\
&= \overline{W}_0^{1/2}\left(-\frac{1-r}{r}U_\tau + \sum\nolimits_{i=\tau+1}^{t}\left(\frac{1}{r} - \frac{1-r}{r}\right)U_i + \frac{1-r}{r}U_{\tau+1}\right) \\
&= \overline{W}_0^{1/2}\left(\sum\nolimits_{i=\tau+1}^{t} U_i + \frac{1-r}{r}\left(U_{\tau+1} - U_\tau\right)\right).
\end{aligned}$$

Hence,

$$\widehat{\delta} = \overline{W}_0^{1/2}\left(\frac{1}{t-\tau}\sum\nolimits_{i=\tau+1}^{t} U_i + \frac{1-r}{r}\frac{U_{\tau+1}-U_\tau}{t-\tau}\right).$$

Evaluating $l\left(\delta, \tau \,\middle|\, \mathbf{U}^*, \overline{Z}_0, \overline{W}_0, r, h, t\right)$ at $\widehat{\delta}$, we have

$$\begin{aligned}
l\left(\widehat{\delta}, \tau \,\middle|\, \mathbf{U}^*, \overline{Z}_0, \overline{W}_0, r, h, t\right) &= -\frac{t}{2}\ln(2\pi) - \frac{t}{2}\ln(r) - \frac{1}{2}\mathbf{G}^{\mathbf{T}}\mathbf{G} - \frac{1}{2}\mathbf{H}^{\mathbf{T}}\mathbf{H} \\
&\quad + \overline{W}_0\left(\frac{1}{t-\tau}\sum\nolimits_{i=\tau+1}^{t} U_i + \frac{1-r}{r}\frac{U_{\tau+1}-U_\tau}{t-\tau}\right)^2 (t-\tau) \\
&\quad - \frac{1}{2}\overline{W}_0\left(\frac{1}{t-\tau}\sum\nolimits_{i=\tau+1}^{t} U_i + \frac{1-r}{r}\frac{U_{\tau+1}-U_\tau}{t-\tau}\right)^2 (t-\tau) \\
&= -\frac{t}{2}\ln(2\pi) - \frac{t}{2}\ln(r) - \frac{1}{2}\mathbf{G}^{\mathbf{T}}\mathbf{G} - \frac{1}{2}\mathbf{H}^{\mathbf{T}}\mathbf{H} \\
&\quad + \frac{1}{2}\overline{W}_0\left(\overline{U}_{t-\tau} + \frac{1-r}{r}\frac{U_{\tau+1}-U_\tau}{t-\tau}\right)^2 (t-\tau).
\end{aligned}$$

It follows that the maximum likelihood prediction for τ is the observed value of

$$(t-\tau)\,\overline{W}_0\left(\overline{U}_{t-\tau} + \frac{1-r}{r}\frac{U_{\tau+1}-U_\tau}{t-\tau}\right)^2$$

for $\tau = 0, 1, \ldots, t-1$. Note that the method depends on the chart parameter r. Note that if we set $\overline{W}_0 = 1$, we have the method for prediction of τ when the in-control process parameters are known.

In the following table, we have the change point predictions when in-control parameters are estimated using simulation based on $10,000$ simulated values for $\tau = 20, 40, 60$ and $\delta = 0.1, \ldots, 1.0$. The values for m and n were set to 4 and 6, respectively. Note that in all cases the predicted values of τ fell between τ and $\tau + 1$. Also in a comparison with the Shewhart $\overline{X}$ chart with estimated in-control parameters, the predicted values are a bit larger with more variability.

	$\tau = 20$			$\tau = 40$			$\tau = 60$	
δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$	δ	$\overline{\tau}$	$s_{\overline{\tau}}$
0.1	20.528	0.499	0.1	40.507	0.500	0.1	60.528	0.499
0.2	20.499	0.500	0.2	40.510	0.500	0.2	60.518	0.500
0.3	20.502	0.501	0.3	40.506	0.500	0.3	60.516	0.500
0.4	20.513	0.500	0.4	40.496	0.500	0.4	60.501	0.500
0.5	20.516	0.500	0.5	40.498	0.500	0.5	60.510	0.500
0.6	20.520	0.499	0.6	40.502	0.500	0.6	60.513	0.501
0.7	20.496	0.500	0.7	40.505	0.500	0.7	60.498	0.500
0.8	20.498	0.500	0.8	40.499	0.500	0.8	60.502	0.500
0.9	20.509	0.500	0.9	40.505	0.500	0.9	60.490	0.500
1.0	20.504	0.500	1.0	40.524	0.499	1.0	60.493	0.500

Table 6.4: In-control Estimated Parameters Case, ARL_0 equal to 200

6.4 Conclusion

A method to detect the change point based on the likelihood function was given both when the in-control process parameters are known and when they are estimated. Change point methods were given for the Shewhart $\overline{X}$ and the EWMA $\overline{X}$ charts. Some comparisons were given.

CHAPTER 7

CONCLUSIONS

7.1 General Conclusions

We have studied the exponentially weighted moving average (EWMA) $\overline{X}$ chart and the ad hoc EWMA$^{(k)}$ $\overline{X}$ charts. Integral equation were given that a useful in studying the the run length performance of the EWMA $\overline{X}$ chart. Simulation results were presented to showing that an EWMA $\overline{X}$ chart could be designed that outperforms the EWMA$^{(2)}$ $\overline{X}$ charts. We presented a chapter on various non-normal families of distribution that would be used to study the robustness of the charts as well as other statistical methods.

A method is given for studying the change point is a production process with respect to a step-change in the process mean. Simulation was used to study compare these methods for the Shewhart $\overline{X}$ chart and the EWMA $\overline{X}$ chart.

7.2 Areas for Further Research

We are interested in studying the robustness of the EWMA $\overline{X}$ and EWMA$^{(k)}$ $\overline{X}$ charts.

Rigdon using simulation that the optimal double EWMA $\overline{X}$ chart is the optimal EWMA $\overline{X}$ chart. Our interest is to use simulation to see if the optimal triple EWMA $\overline{X}$ chart is an optimal EWMA $\overline{X}$ chart. The difficulties in performing these simulations is determining the formulas that given the variance of the triple EWMA $\overline{X}$ chart. Further we conjecture that the double EWMA $\overline{X}$ chart outperforms the triple EWMA $\overline{X}$ chart and the triple EWMA $\overline{X}$ chart outperforms the quadruple EWMA $\overline{X}$ chart.

It is often recommended that the practitioner monitor for both a change in the process mean and process variance or standard deviation. We are interest in studing the change point with respect to a change in the process standard deviation as well as the process mean.

We are interested in developing integral equations that can be used to study the performance of the EWMA$^{(k)}$ $\overline{X}$ charts.

REFERENCES

[1] Roberts, S.W. (1959), ”Control Chart Tests Based on Geometric Moving Averages,” Technometrics 1(3), 239-250.

[2] Champ, C.W., Rigdon, S.E., and Scharnagl, K.A. (2001), “Method for Deriving Integral Equations Useful in Control Chart Performance Analysis,” Nonlinear Analysis: Theory, Methods, and Applications 47(3), 2089-2101

[3] Duncan, A.J. (1989), Quality Control and Industrial Statistics, Fourth Edition, McGraw-Hill: New York.

[4] Shewhart, W. A. (1931). Economic Control of Quality of Manufactured Product. Macmillan And Co Ltd, London

[5] Alevizakos, V., Chatterjee, K., and Koukouvinos, C. (2021a) “The Triple Exponentially Weighted Moving Average Control Chart,” Quality Technology Quantitative Management 18(3), 326-354. DOI: 10.1080/16843703.2020.1809063.

[6] Alevizakos, V., Chatterjee, K., and Koukouvinos, C. (2021b) “The Quadru- ple Exponentially Weighted Moving Average Control Chart,” Quality Technology Quantitative Management 19(1), 50-73. DOI: 10.1080/16843703.2021.1989141.

[7] Burr, I. W. (1976). Statistical Quality Control Methods, Volume 16. CRC Press.

[8] Champ, C.W. Woodall, W.H., and Mohsen, H.A. (1991), “A Generalized Quality Control Procedure,” Statistics and Probability Letters 11, 211-218.

[9] Chan, L. K., Hapuarachchi, K., and Macpherson, B. (1988). Robustness of Mean e (x) and r Charts. IEEE Transactions on Reliability, 37(1):117–123.

[10] Jones,Champ and Rigdon. (2001), “Performance Analysis of Exponentially Weighted Moving Average Charts When Parameters Are Estimated,” Technometrics 43, 156-167.

[11] Knoth, S., Saleh, N.A., Mahmond, M.A., Woodall, W.H. and Tercero-Gomez, V.G. (2021), "A Critique of a Variety of "Memory-Based" Process Monitoring Methods," Journal of Quality Technology 55, 18-42.

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

[13] Mahmoud, M.A. and Woodall, W.H. (2010), "An Evaluation of the Double Exponentially Weighted Moving Average Control Chart," Communications in Statistics - Simulation and Computation 39(5), 933-949.

[14] Montgomery, D. C. (2009). Statistical Quality Control, Volume 7. Wiley New York.

[15] Montgomery, D. C. (2020). Introduction to Statistical Quality Control. John Wiley & Sons.

[16] Schilling, E. G. and Nelson, P. R. (1976). The Effect of Nonnormality on The Control Limits of x Charts. Journal of Quality Technology, 8(4):183–188.

[17] Shamma, S.E. and Shamma, A.K. (1992), "Development and Evaluation of Control Charts Using Double Exponentially Weighted Moving Averages," International Journal of Quality Reliability Management Vol. 9(6), 18-25. https://doi.org/10.1108/02656719210018570

[18] Shewhart, W. A. (1931). Economic Control of Quality of Manufactured Product. Macmillan And Co Ltd, London.

[19] Woodall, W.H. (1983), "The Distribution of the Run Length of One-Sided CUSUM Procedures for Continuous Random Variables," Technometrics 25(3),295-301.

[20] Yourstone, S. A. and Zimmer, W. J. (1992). Non-normality and the Design of Control Charts for Averages. Decision Sciences, 23(5):1099–1113.

[21] Hull, W.E. and Dobell, A.R. (1962), "Random Number Generators," SIAM Review 4(3), 230-254