OPTIMAL MONITORING METHODS FOR UNIVARIATE AND MULTIVARIATE EWMA CONTROL CHARTS

By

ICK HUH, B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2014) McMaster University

(Statistics) Hamilton, Ontario

TITLE: Optimal Monitoring Methods for Univariate and Multivariate EWMA Control Charts

AUTHOR: Ick Huh

(McMaster University, Canada)

SUPERVISOR: Professor Román Viveros-Aguilera

Professor N. Balakrishnan

NUMBER OF PAGES: xvi, 134

Abstract

Due to the rapid development of technology, quality control charts have attracted more attention from manufacturing industries in order to monitor quality characteristics of interest more effectively. Among many control charts, my research work has focused on the multivariate exponentially weighted moving average (MEWMA) and the univariate exponentially weighted moving average (EWMA) control charts by using the Markov chain method. The performance of the chart is measured by the optimal average run length (ARL). My Ph.D. thesis is composed of the following three contributions.

My first research work is about differential smoothing. The MEWMA control chart proposed by Lowry et al. (1992) has become one of the most widely used charts to monitor multivariate processes. Its simplicity, combined with its high sensitivity to small and moderate process mean jumps, is at the core of its appeal. Lowry et al. (1992) advocated equal smoothing of each quality variable unless there is an a priori reason to weigh quality characteristics differently. However, one may have situations where differential smoothing may be justified. For instance: (a) departures in process mean may be different across quality variables, (b) some variables may evolve over time at a much different pace than other variables, and (c) the level of correlation between variables could vary substantially. For these reasons, I focus on and assess the performance of the differentially smoothed MEWMA chart. The case of two quality variables (BEWMA) is discussed in detail. A bivariate Markov chain method that uses conditional distributions is developed for average run length (ARL) calculations. The proposed chart is shown to perform at least as well as Lowry et al. (1992)'s chart, and noticeably better in most other mean jump directions. Comparisons with the recently introduced double-smoothed BEWMA chart and the univariate charts for the independent case show that the proposed differentially smoothed BEWMA chart has superior performance. This part of our work was published recently in *Journal of Quality Technology*

(**45**(4), pp. 377–393; 2013).

My second research work is about monitoring skewed multivariate processes. Recently, Xie et al. (2011) studied monitoring bivariate exponential quality measurements using the standard MEWMA chart originally developed to monitor multivariate normal quality data. The focus of my work is on situations where, marginally, the quality measurements may follow not only exponential distributions but also other skewed distributions such as Gamma or Weibull, in any combination. The joint distribution is specified using the Gumbel copula function thus allowing for varying degrees of correlation among the quality measurements. In addition to the standard MEWMA chart, charts based on the largest or smallest of the measurements and on the joint cumulative distribution function or the joint survivor function, are studied in detail. The focus is on the case of two quality measurements, i.e., on skewed bivariate processes. The proposed charts avoid an undesirable feature encountered by Xie et al. (2011) for the standard MEWMA chart where in some cases the off-target average run length turns out to be larger than the on-target one. Using the optimal average run length, our extensive numerical results show that the proposed methods provide an overall good detection performance in most directions. Simulations were performed to obtain the optimal ARL results but the Markov chain method using the empirical CDF of the statistics involved verified the accuracy of the ARL results. In addition, an examination of the effect of correlation on chart performance was undertaken numerically. The methods are easily extendable to more than two variables.

Final study is about a new ARL criterion for univariate processes studied in detail in this thesis. The traditional ARL is calculated assuming a given fixed process mean jump and a given time point where the jump occurs, usually taken to be from the very beginning in most chart performance studies. However, Ryu et al. (2010) demonstrated that the assumption of a fixed mean shift might lead to poor performance of control charts when the actual size of the mean shift is significantly different and therefore suggested a new ARL-based performance measure, called expected weighted run length (EWRL), by assuming that

the size of the mean shift is not specified but rather it follows a probability distribution. The EWRL becomes the expected value of the weighted ordinary ARL with respect to this distribution. My methods generalize this criterion by allowing the time at which the mean shift occurs to also vary according to a probability distribution. This leads to a joint distribution for the size of the mean shift and the time the shift takes place, then the EWRL is calculated as the weighted expected value with respect to this joint distribution. The benefit of the generalized EWRL is that one can assess the performance of control charts more realistically when the process starts on-target and then the mean shift occurs at some later random time. Moreover, I also propose the effective EWRL, which measures the number of additional process runs that on average are needed to detect a jump in the mean after it happens. I evaluate several well-known univariate control charts based on their EWRL and effective EWRL performance. The numerical results show that the choice of control chart depends on the additional information on the transition point of the mean shift. The methods can readily be extended to other control charts, including multivariate charts.

Key words: *Average Run Length; Bivariate Exponentially Weighted Moving Average Control Chart; Differential Smoothing Scheme; Empirical Distribution Function; Expected Weighted Run Length; Markov Chain Method; Non-centrality Parameter; Optimal Average Run Length; Single Smoothing Scheme; Smoothing Parameter; Survival Gumbel Copula; Transition Point.*

Acknowledgement

First and foremost, I would like to express the deepest gratitude to my doctoral supervisors, Professor Román Viveros-Aguilera and Professor Narayanaswamy Balakrishnan, for their wisdom, immense knowledge, encouragement and helpful suggestions which have been useful throughout my thesis work. Without their thoughtful guidance, this work would not have been successful. I also would like to acknowledge valuable advice from other member of my supervisory committee, Professor Aaron Childs. I thank the Department of Mathematics and Statistics at McMaster University for providing financial support. I thank Compute Canada for providing computing resources which made my computational work possible. I am grateful to all my friends for their support. Last, but not the least, I would like to thank my family members, especially my wife, Vivian, for supporting and encouraging me to pursue this degree. Without her support, none of this would have been possible.

Publications from the thesis

1. Huh, I., Viveros, R., and Balakrishnan, N. (2013). "Differential Smoothing in the Bivariate Exponentially Weighted Moving Average Chart." *Journal of Quality Technology.* **45**(4), 377-393.

2. Huh, I., Viveros, R., and Balakrishnan, N. "EWMA Chart for Monitoring Bivariate Skewed Data." *Under Preparation.*

3. Huh, I., Viveros, R., and Balakrishnan, N. "Designing Optimal Charts with Prior Information." *Under Preparation.*

List of Abbreviations and Acronyms

ARL	Average Run Length
ARL_{opt}	Optimal Average Run Length
ARL_{opt}^{ef}	Optimal Effective Average Run Length
ARL_0	In-Control Average Run Length
ARL_1	Out-of-ControlAverage Run Length
BEWMA	Bivariate Exponentially Weighted Moving Average
CDF	Cumulative Distribution Function
CUSUM	Cumulative Sum
dBEWMA	Double Bivariate Exponentially Weighted Moving Average
dMEWMA	Double Multivariate Exponentially Weighted Moving Average
D-D	Downward-Downward
D-U	Downward-Upward
EARL	Expected Average Run Length
EARL^{ef}	Expected Effective Average Run Length
ERARL	Expected Relative Average Run Length
ERARL_{opt}	Optimal Expected Relative Average Run Length
$\mathsf{ERARL}_{opt}^{ef}$	Optimal Expected Relative Effective Average Run Length
EWMA	Exponentially Weighted Moving Average
EWRL	Exponentially Weighted Run Length
FIR	Fast Initial Response
HS	Head Start
IC	In-Control
IEWMA	Improved Exponentially Weighted Moving Average
LCL	Lower Control Limit
MAX-MIN	Maximum-Minimum
MEWMA	Multivariate Exponentially Weighted Moving Average

OOC	Out-of-Control
REWMA	Reset Exponentially Weighted Moving Average
SPC	Statistical Process Control
UCL	Upper Control Limit
U-D	Upward-Downward
U-U	Upward-Upward

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Quality Control Charts

1.1 Control Charts

Control charts are some of the most useful tools for monitoring and improving the quality of products in statistical process control (SPC). They are often displayed graphically to visualize how a process changes over time. Depending on the number of characteristics of interest, they are classified into two basic types of control charts: a univariate and a multivariate control charts. For illustration purpose, let us look at the Shewhart $\overline{X}$ chart which is arguably the most popular univariate control chart (Figure 1.1). In general, as Figure 1.1 presents, control charts contain the control limit(s) and the centre line representing the average value of the quality characteristic corresponding to the in-control state. The points represent a statistic of the measurements of a quality characteristic in the samples. Note that in Phase I analysis, based on data gathered when the the process is in-control, the process parameters are estimated and the control limits are determined. This stage is also referred to as chart calibration. In Phase II analysis, the chart is put in place in usual process operation and used to detect out-of-control process excursions.

As long as the points are observed within the control limits, the process is assumed to be in-control or stable. If they are observed outside of the limits, the process is out-of-control. When the process is out-of-control, it is required to take action to remove the cause of the out-of-control state. For the definitions of the in-control and out-of-control states, refer to Sections 1.4.1 and 1.4.2.

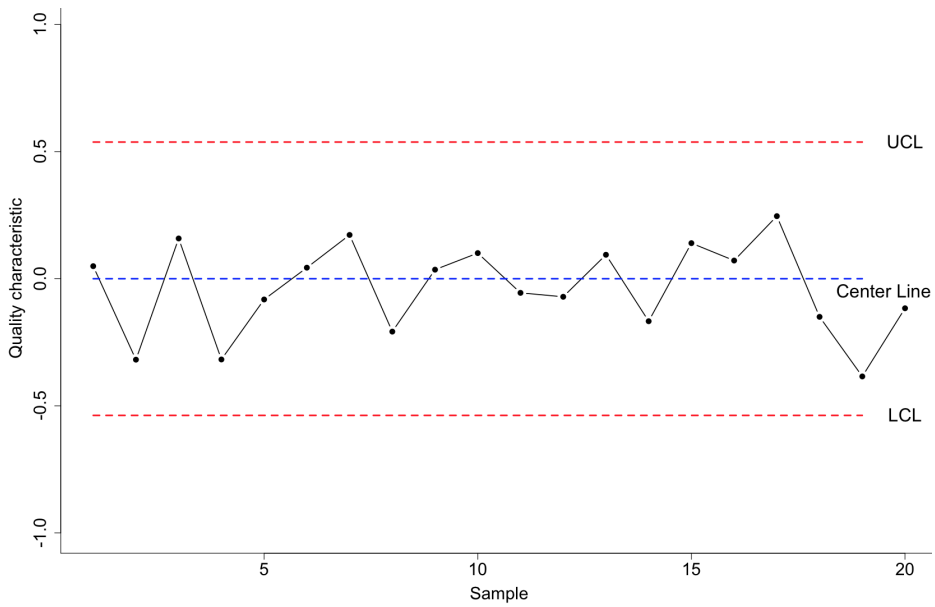


Figure 1.1: Shewhart Control Chart.

1.2 Univariate Control Charts

As the name implies, the univariate control charts are used to monitor a single quality characteristic. The univariate control charts have been popular for monitoring business or industrial process data due to their simplicity. Among many univariate control charts, Shewhart chart, the exponentially weighted moving average (EWMA) chart and the cumulative sum (CUSUM) chart are explained in detail below. Denote the observations of the monitoring process by $X_1, X_2, \ldots,$ which are assumed to be independent.

1.2.1 Shewhart Control Chart

In monitoring one characteristic of interest, the Shewhart charts have been very useful in detecting large mean shifts. Among many Shewhart-type control charts, the Shewhart $\overline{X}$ charts are widely used in practice. Suppose that the quality characteristic of interest follows a normal distribution with mean μ and variance σ^2. Assume that a random sample of size n is taken at each sampling point t and $X_{t1}, X_{t2}, \ldots, X_{tn}$ is a sample of size n. Then, the average of this sample is

$$\overline{X_t} = \frac{X_{t1} + X_{t2} \cdots + X_{tn}}{n}.$$

Also, $\overline{X}_t$ follows a normal distribution with mean μ and variance $\frac{\sigma^2}{n}$. Moreover, the upper control limit (UCL) and the lower control limit (LCL) are $\mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and $\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$, respectively, where $z_{\alpha/2}$ is a $\alpha/2$ quantile of the standard normal distribution. For 95.44% confidence, $z_{\alpha/2} = 2$ while $z_{\alpha/2} = 3$ for 99.74%. The control chart signals an alarm if an observation falls outside the control limits. Note that Shewhart control charts could be used in both Phase I analysis and Phase II analysis but the charts are more useful in Phase I analysis due to their ability to detect large shifts (Montgomery, 2009, pp. 198).

1.2.2 Exponentially Weighted Moving Average (EWMA) Chart

It is well-known that the exponentially weighted moving average (EWMA) charts (Roberts 1959) provide quicker detection of small and moderate mean shifts than the Shewhart control charts. Suppose that we have one characteristic of interest and observe $X_1, X_2, \ldots$. Then, the EWMA chart (Roberts 1959) is given by

$$Z_t = rX_t + (1-r)Z_{t-1}, \quad t = 1, 2, \ldots, \tag{1.1}$$

where r is the smoothing parameter such that $0 < r \leq 1$. The initial value of EWMA statistic is $Z_0 = E[X]$. The smoothing parameter balances present data and historical data

by assigning the weight. There is no generally accepted guidance for choosing the smoothing parameter, but Montgomery (2009) recommends $0.05 \leq r \leq 0.25$ with $r = 0.1$ being the default value. Note that when $r = 1$, the chart becomes the Shewhart $\overline{X}$ control chart. If $X_1, X_2 \ldots$ are independent and identically distributed following a normal distribution $N(\mu_0, \sigma^2)$, then the mean of Z_t is μ_0 and the variance is

$$\frac{r}{2-r}\left[1-(1-r)^{2t}\right]\sigma^2. \tag{1.2}$$

Note that, as t goes to infinity, Equation (1.2) reduces to

$$\frac{r}{2-r}\sigma^2. \tag{1.3}$$

Asymptotically, the control limits of the EWMA chart are placed at

$$\mu_0 \pm L\sqrt{\frac{r}{2-r}}\sigma,$$

where L is a parameter. The chart sends a signal when the statistic Z_t falls outside of the control limits.

1.2.3 Cumulative Sum Control (CUSUM) Chart

Suppose that $X_1, X_2, \ldots$ follow the standard normal distribution. The one-sided upper CUSUM statistic (Page 1954) is defined as

$$C_t^+ = \max\left[0, X_t - k + C_{t-1}^+\right], \quad t = 1, 2, \ldots,$$

where $E[X_t] = 0$, k is the reference value and $C_0 = 0$. If C_t^+ exceeds a control limit h, an alarm is triggered. The one-sided lower CUSUM statistic is given by

$$C_t^- = \max\left[0, -X_t - k + C_{t-1}^-\right], \quad t = 1, 2, \ldots$$

With respect to the role of the reference value k, it is somewhat similar to that of the smoothing parameter in the EWMA control chart. Note that the CUSUM control charts can be optimally designed by setting k as half of a set mean shift. The CUSUM charts are also more effective than the Shewhart charts for detecting small mean shifts.

1.3 Multivariate Control Charts

In recent years, the importance of multivariate control charts has increased because more quality features are measured in mass production than ever before. Applying separate univariate control charts to monitor each individual variable is a possible approach. However, it is clear that managing separate univariate charts is not an ideal approach and it might lead to erroneous conclusions. Thus, when quality characteristics are cross-correlated, it would be more efficient to maintain one multivariate control chart rather than use separate univariate control charts. Suppose that we have p quality characteristics of interest and the $p \times 1$ random vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots$ denote the observations vectors over time. It is assumed that $\mathbf{X}_1, \mathbf{X}_2, \ldots$ are independent random vectors with mean vectors $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots$ respectively. Furthermore, observations are assumed to have a common covariance matrix $\Sigma_{\mathbf{X}}$ when the process is in-control. Several multivariate control charts have been suggested for monitoring changes in the quality characteristics. In the following subsections, some of the most common multivariate control charts are reviewed.

1.3.1 Hotelling χ^2 Control Chart

Probably, the Hotelling χ^2 control chart is considered to be the most popular multivariate control chart for monitoring several quality characteristics. This is a direct analog of the univariate Shewhart $\overline{X}$ chart. Suppose that $\mathbf{X}_t$ has a multivariate normal distribution. Then, Hotelling χ^2 chart (Hotelling 1947) gives an out-of-control signal as soon as

$$\chi_t^2 = (\mathbf{X}_t - \boldsymbol{\mu_0})'\boldsymbol{\Sigma}_{\mathbf{X}}^{-1}(\mathbf{X}_t - \boldsymbol{\mu_0}) > h, \quad t = 1, 2, \ldots,$$

where h is a specified control limit and it is obtained from a chi-squared distribution with p degrees of freedom. Moreover, if the process mean changes from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$ and else remaining the same, χ_t^2 follows a non-central chi-squared distribution with the non-centrality parameter c given by

$$c = (\boldsymbol{\mu}_1 - \boldsymbol{\mu_0})'\boldsymbol{\Sigma}_{\mathbf{X}}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu_0}).$$

It is generally accepted that the Hotelling χ^2 control chart is not effective in detecting small and moderate mean shifts. Moreover, as the number of process variables increases, traditional multivariate control charts such as the Hotelling's χ^2 control chart lose efficiency with respect to detection power.

1.3.2 MEWMA Control Chart

In the multivariate case, the multivariate exponentially weighted moving average (MEWMA) chart (Lowry et al. 1992) is a natural extension of the EWMA chart (Equation (1.1)). The MEWMA chart is based on the values

$$\mathbf{Z}_t = \mathbf{R}(\mathbf{X}_t - \boldsymbol{\mu}_0) + (\mathbf{I} - \mathbf{R})\mathbf{Z}_{t-1}, \quad t = 1, 2, \ldots, \tag{1.4}$$

where $\mathbf{Z}_0 = \mathbf{0}$ and $\mathbf{R} = \text{diag}(r_1, r_2, \ldots, r_p),\ 0 < r_j \leq 1,\ j = 1, 2, \ldots, p$. The chart issues an alarm as soon as

$$Q_t = \mathbf{Z}_t' \Sigma_{\mathbf{Z}_t}^{-1} \mathbf{Z}_t > h, \tag{1.5}$$

where h is chosen to obtain a specified in-control ARL and $\mathbf{R}$ is the smoothing parameter matrix. In practice, it is generally assumed that $r_1 = r_2 = \cdots = r_p = r$ and $\boldsymbol{\mu}_0 = \mathbf{0}$. Then, Equation (1.4) can be rewritten as

$$\mathbf{Z}_t = r\mathbf{X}_t + (1 - r)\mathbf{Z}_{t-1}, \quad t = 1, 2, \ldots. \tag{1.6}$$

By taking an asymptotic approach ($t \to \infty$), the asymptotic covariance matrix of $\mathbf{Z}_t$ is $\Sigma_{\mathbf{Z}} = \{r/(2 - r)\}\, \Sigma_{\mathbf{X}}$, where $\Sigma_{\mathbf{X}}$ is the covariance matrix of $\mathbf{X}_t$. Note that since one single smoothing parameter is used for the MEWMA chart, the performance of a MEWMA chart is a function of the mean shift ($\boldsymbol{\mu}_1$) only through the non-centrality parameter ($c = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}$). For example, suppose that we have two different mean shifts ($\boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$) from the in-control mean vector ($\boldsymbol{\mu}_0 = \mathbf{0}$). Then, as long as we have $\sqrt{\boldsymbol{\mu}_1' \Sigma_{\mathbf{X}}^{-1} \boldsymbol{\mu}_1} = \sqrt{\boldsymbol{\mu}_2' \Sigma_{\mathbf{X}}^{-1} \boldsymbol{\mu}_2}$, we will always have the same performance (i.e., the same ARL). This property of the MEWMA chart is directional invariance. Just as with the EWMA chart, the MEWMA chart is effective for detecting small and moderate size mean shifts. When it comes to detecting small mean shifts, the use of small values for the smoothing parameter increases the power of the control chart. Note that the Hotelling χ^2 control chart is a special case of the the MEWMA chart, obtained when $r = 1$.

1.4 Average Run Length

The average run length (ARL) is one of the most popular measures to quantify the performance of a control chart. In general, the notion of the ARL is the average number of

points plotted before the chart gives a signal. Denote by N the number of process runs until a control chart first signals. Then, N is called the run length for the scheme and the probability distribution of N is the run length distribution. Figure 1.2 shows typical shapes of the distribution of run length. They are the on-target run length distributions for the MEWMA chart with $\text{ARL}_0 = 200$, 400 and 600, respectively. As seen in Figure 1.2, it is a highly skewed distribution (positively skewed) and the mean of the distribution is the ARL. Note that for a Shewhart control chart, the distribution of run length is a geometric distribution. In this case, the ARL can be easily calculated as $\text{ARL} = 1/p$ where p is the probability that any point exceeds the control limits.

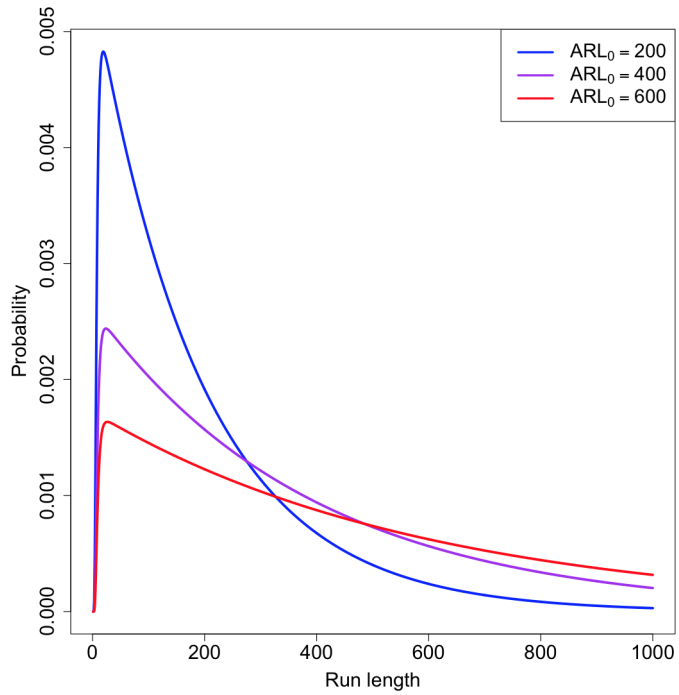


Figure 1.2: The Shape of the Run Length Distribution (obtained by the MEWMA chart with the smoothing parameter, $r = 0.1$)

.

There are two types of ARL: on-target (in-control) ARL which is often denoted by ARL_0, and off-target (out-of-control) ARL which is denoted by ARL_1.

1.4.1 On-target ARL (= $\mathbf{ARL}_0$)

The definition of "in-control (on-target) process" is that a process is operating with only chance causes (or common causes) of variation being present. Chance causes are a natural and inherent part of any production process. Suppose that the process is in-control. Then, the ARL_0 is the average number of points plotted before a false alarm is raised for the first time. For example, if ARL_0 is 200, it is expected that the chart signals a false alarm on average every 200 plotted points despite the fact that the process is stable. Note that it is recommended that a larger ARL_0 be used in order to reduce the frequency of false alarms.

1.4.2 Off-target ARL (= $\mathbf{ARL}_1$)

"Out-of-control (off-target or unstable) process" means that a process is operating in the presence of assignable causes. Machine malfunction, operator mistakes and power surges belong to assignable causes. When the process is out-of-control, the control chart shows that data points do not cluster around the mean value and fall the outside of the control limits. Suppose that the process is out-of-control. Then, ARL_1 is defined as the average number of points required to detect a shift. For example, let us say that ARL_1 is 68.1 for a particular mean shift. That is, on average the chart needs 68.1 samples (runs) to detect the mean shift. Note that in order to detect a shift efficiently, smaller ARL_1 is required. Additionally, in order to optimally design the control chart at a specified mean shift, the optimal ARL_1 is used. For example, when the MEWMA chart is optimized for a particular mean shift, the optimal ARL_1 (= ARL_{opt}) is the minimum value of ARL_1 achieved when the smoothing parameter r is varied. Denote by r_{opt} the value of r yielding the smallest ARL_1.

1.5 Markov Chain Method

Consider a discrete time stochastic process with finite space, $\{X_n, n = 0, 1, 2, \dots\}$. A Markov chain is a sequence of random variables $X_0, X_1, X_2 \dots$, having the property that, given the present, the future and the past states are independent. In other words,

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_1 = x_1, X_0 = x_0) = P(X_n = x_n | X_{n-1} = x_{n-1}). \tag{1.7}$$

Equation (1.7) is the so-called transition probability and the Markov chain approach is often used to approximate the ARL of a control chart. Suppose that Z_t is a control statistic, where $t = 0, 1, 2, \dots$ and the process Z_t is a Markov chain. Then, obtain a discretized Markov chain by dividing the in-control region which is bounded by control limits into many sub-regions. The area outside the control limits represents the absorbing state while each sub-region within the in-control region is called a transient state. An absorbing state is a state which it is impossible to leave, once entered. In other words, a state x is called absorbing if and only if $P(Z_t = x | Z_{t-1} = x) = 1$ and $P(Z_t = y | Z_{t-1} = x) = 0$ for $x \neq y$. Denote the probability of first recurrence to x at the nth step by $f_{xx}^{(n)} = P(Z_n = x, Z_{n-1} \neq x, \dots, Z_1 \neq x | Z_0 = x)$. Then, if $\sum_{n=1}^{\infty} f_{xx}^{(n)} < 1$, a state x is called transient (Karlin and Talyor, 1975, pp. 64). At time t, the process is said to be in-control if Z_t is in a transient state and the process is said to be out-of-control when Z_t is in an absorbing state. By computing all transition probabilities from one transient state to another transient state, the transition probability matrix can be constructed. Then, the ARL is given by

$$\text{ARL} \doteq \mathbf{s}'(\mathbf{I} - \mathbf{P})^{-1}\mathbf{1},$$

where $\mathbf{s}$ is the initial probability vector for the states, $\mathbf{1}$ is a column vector of 1s, $\mathbf{I}$ is the identity matrix and $\mathbf{P}$ is the transition probability matrix.

1.5.1 Markov Chain Method for the EWMA Chart

Lucas and Saccucci (1990) set up a Markov chain method for the EWMA chart by discretizing the control statistic. Divide the in-control region (the interval between the control limits LCL and UCL) into $2m+1$ subintervals and let S_j represent j^{th} subinterval, where $j = 1, 2, \ldots, 2m+1$. For the approximation, the midpoint of the jth subinterval, s_j is taken representative value for S_j and we define δ as the half size of each subinterval, that is $\delta = \dfrac{\text{UCL} - \text{LCL}}{2(2m+1)}$. If $s_j - \delta < Z_t < s_j + \delta$ for $j = 1, 2, \ldots, 2m+1$, then the process is said to be a transient state; otherwise, it is said to be in the absorbing state. Thus, the in-control transition probability from one transient state S_j to another transient state S_k can be written as follows:

$$\begin{aligned} p_{kj} &= P\left[Z_t \in S_k \mid Z_{t-1} \in S_j\right] \\ &\doteq P\left[s_k - \delta < rX_t + (1-r)Z_{t-1} < s_k + \delta \mid Z_{t-1} = s_j\right] \\ &= P\left[\{s_k - \delta - (1-r)s_j\}/r < X_t < \{s_k + \delta - (1-r)s_j\}/r\,\right], \qquad (1.8) \\ &\quad \text{where } j, k = 1, 2, \ldots, 2m+1. \end{aligned}$$

If $X_t \sim N(\mu_0, \sigma^2)$, then Equation (1.8) can be rewritten as follows:

$$\begin{aligned} p_{kj} &= P\left[Z_t \in S_k \mid Z_{t-1} \in S_j\right] \\ &\doteq \Phi\left[\{(s_k + \delta) - (1-r)s_j - r\mu_0\}/(r\sigma)\right] - \Phi\left[\{(s_k - \delta) - (1-r)s_j - r\mu_0\}/(r\sigma)\right], \\ &\quad \text{where } j, k = 1, 2, \ldots, 2m+1, \text{ and } \Phi \text{ is the cdf of the standard normal distribution.} \end{aligned}$$

Moreover, when the in-control process mean shifts from μ_0 to μ_1, the out-of-control transition probability from one transient state S_j to another transient state S_k can be obtained

as follows:

$$
\begin{aligned}
p_{kj} &= P\left[Z_t \in S_k \mid Z_{t-1} \in S_j\right] \\
&\doteq P\left[s_k - \delta < rX_t + (1-r)Z_{t-1} < s_k + \delta \mid Z_{t-1} = s_j\right] \\
&= \Phi\left[\{(s_k+\delta) - (1-r)s_j - r\mu_1\}/(r\sigma)\right] - \Phi\left[\{(s_k-\delta) - (1-r)s_j - r\mu_1\}/(r\sigma)\right], \\
&\quad \text{where } j,k = 1,2,\ldots,2m+1.
\end{aligned}
$$

The transition probability matrix becomes

$$\mathbf{P} = (p_{kj})_{k=1}^{2m+1}{}_{j=1}^{2m+1}.$$

1.5.2 Markov Chain Method for the MEWMA Chart

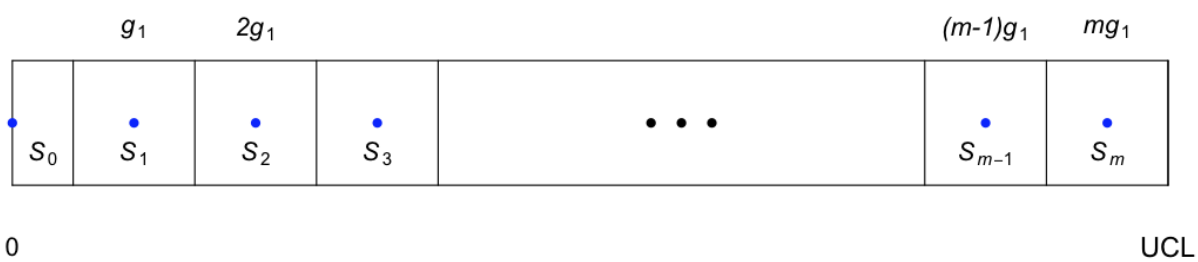


Figure 1.3: An Illustration of Partitioning of the In-Control Region of a MEWMA Chart.

Prabhu and Runger (1996) used the Markov chain method to calculate the run length performance of a MEWMA control chart. For the in-control performance, Figure 1.3 illustrates the transient states when the process is in-control by discretizing the chart statistic (Equation (1.6)). They used $q_t = ||\mathbf{Z}_t||$ as the statistic, which is a measure of distance in p dimensional space, where $||\mathbf{Z}_t|| = \sqrt{\mathbf{Z}_t'\mathbf{Z}_t}$. Thus, $Q_t > h$ (Equation (1.5)) is equivalent to $q_t = ||\mathbf{Z}_t|| > \sqrt{\dfrac{rh}{2-r}}$. Consequently, we set UCL $= \sqrt{\dfrac{rh}{2-r}}$ and the in-control region is $[0, \text{UCL}]$. Dividing the in-control region $[0, \text{UCL}]$ into $m+1$ subintervals, m of them having the same length g_1 and the very first subinterval having the length $g_1/2$, where

$g_1 = (2\text{UCL})/(2m+1)$. $S_i, i = 0, 1, 2, \ldots, m$, represent transient states and the point in each subinterval represents the midpoint of the interval. The distribution of q_t can be approximated as follows. For $j = 0, 1, 2, \ldots, m$ and k not equal to 0, the transition probability of q_t from state j to state k is given by

$$
\begin{aligned}
& P\left[q_t \in S_k \mid q_{t-1} \in S_j\right] \\
= \; & P\left[(k-0.5)g_1 < ||r\mathbf{X}_t + (1-r)\mathbf{Z}_{t-1}|| < (k+0.5)g_1 \mid ||\mathbf{Z}_{t-1}|| \in S_j\ \right] \\
\doteq \; & P\left[(k-0.5)g_1 < ||r\mathbf{X}_t + (1-r)\mathbf{Z}_{t-1}|| < (k+0.5)g_1 |\ q_{t-1} = jg_1\right] \\
& \text{(Note that } \mathbf{Z}_{t-1} \text{ is a spherical distribution (see Appendix A.1) and the distribution} \\
& \text{of } \mathbf{Z}_{t-1}, \text{given } ||\mathbf{Z}_{t-1}||, ||\mathbf{Z}_{t-2}||, \ldots, ||\mathbf{Z}_1||, \text{ is uniform } (\mathbf{U}) \text{ on the } p \text{ dimensional sphere} \\
& \text{of radius } ||\mathbf{Z}_{t-1}||). \\
= \; & P\left[(k-0.5)g_1 < ||r\mathbf{X}_t + (1-r)jg_1\mathbf{U}|| < (k+0.5)g_1\right] \\
= \; & P\left[(k-0.5)g_1/r < ||\mathbf{X}_t + (1-r)jg_1\boldsymbol{e}/r|| < (k+0.5)g_1/r\right] \\
= \; & P\left[(k-0.5)^2 g_1^2/r^2 < \chi^2(p,c) < (k+0.5)^2 g_1^2/r^2\right], \qquad (1.9) \\
& \text{where } \boldsymbol{e}' = (1,0,0,\ldots,0), c = [(1-r)jg_1/r]^2 \text{and } \chi^2(p,c) \text{ is a non-central chi-squared} \\
& \text{distribution with } p \text{ degrees of freedom and non-centrality parameter } c.
\end{aligned}
$$

For $k = 0$, we have

$$
\begin{aligned}
& P\left[q_t \in S_0 \mid q_{t-1} \in S_j\right] \\
\doteq \; & P\left[\chi^2(p,c) < (0.5)^2 g_1^2/r^2\right]. \qquad (1.10)
\end{aligned}
$$

For the out-of-control performance, Prabhu and Runger (1996) used a bivariate Markov

chain method. Let us assume that

$$E[\mathbf{X}] = \begin{cases} \boldsymbol{\mu_0} & \text{when the process is in-control} \\ \boldsymbol{\mu_1} & \text{when the process is out-of-control} \end{cases}$$

and $Var[\mathbf{X}] = \Sigma_{\mathbf{X}}$. If we consider the variable $\Sigma_{\mathbf{X}}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}_0)$, then we obtain

$$E\left[\Sigma_{\mathbf{X}}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}_0)\right] = \begin{cases} \Sigma_{\mathbf{X}}^{-1/2} E\left[\mathbf{X} - \boldsymbol{\mu_0}\right] = \mathbf{0} & \text{when the process is in-control} \\ \Sigma_{\mathbf{X}}^{-1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu_0}) & \text{when the process is out-of-control.} \end{cases}$$

At the same time, we have $Var\left[\Sigma_{\mathbf{X}}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}_0)\right] = \Sigma_{\mathbf{X}}^{-1/2}\Sigma_{\mathbf{X}}\left(\Sigma_{\mathbf{X}}^{-1/2}\right)' = \mathbf{I}$. As a result of that, the non-centrality parameter of the variable $\Sigma_{\mathbf{X}}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}_0)$ is

$$\begin{aligned} c &= \left(\Sigma_{\mathbf{X}}^{-1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu_0}) - \mathbf{0}\right)' (\mathbf{I})^{-1} \left(\Sigma_{\mathbf{X}}^{-1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu_0}) - \mathbf{0}\right) \\ &= (\boldsymbol{\mu}_1 - \boldsymbol{\mu_0})' \Sigma_{\mathbf{X}}^{-1/2}\Sigma_{\mathbf{X}}^{-1/2}(\boldsymbol{\mu}_1 - \boldsymbol{\mu_0}) \\ &= (\boldsymbol{\mu}_1 - \boldsymbol{\mu_0})' \Sigma_{\mathbf{X}}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu_0}). \end{aligned}$$

Thus, the transformation shows that the performance of a MEWMA chart is a function of $\boldsymbol{\mu}_1$ only through the non-centrality parameter (Lowry et al. 1992). Moreover, we can assume that $\mathbf{X}$ has mean zero and an identity covariance matrix without loss of generality.

Since the ARL is a function of the non-centrality parameter, when the process mean shifts to $\boldsymbol{\mu}_1$ from the zero vector ($\boldsymbol{\mu}_0 = \mathbf{0}$), the mean vector $\boldsymbol{\mu}_1$ can be taken to be $\boldsymbol{\mu}_1 = \delta \boldsymbol{e}$, where $\delta = ||\boldsymbol{\mu}_1||$, without loss of generality. Thus, the chart statistic $\mathbf{Z}_t$ can break up into two parts:

$$q_t = ||\mathbf{Z}_t|| = (Z_{t1}^2 + \mathbf{Z}_{t2}'\mathbf{Z}_{t2})^{1/2},$$

Z_{t1} is a one-dimensional random variable with nonzero mean δ and $\mathbf{Z}_{t2}$ is a $p-1$ dimensional

random vector with zero mean vector. Note that Z_{t1} and $\mathbf{Z}_{t2}$ are independent. Figure 1.4 illustrates the two-dimensional graph of $(Z_{t1}, ||\mathbf{Z}_{t2}||)$ with the axes Z_{t1} and $||\mathbf{Z}_{t2}||$. For the approximation of Z_{t1}, as Figure 1.4 indicates, there are $2m_2 + 1$ transient states between between -UCL and UCL and the width of each state, g_2 is $2\text{UCL}/(2m_2 + 1)$. Thus, the transition probability of Z_{t1} from transient state i_x to transient state j_x is as follows. For i_x, $j_x = 1, 2, \ldots, 2m_2 + 1$,

$$\begin{aligned}
& P\left[Z_{t1} \text{ in state } j_x \mid Z_{t1-1} \text{ in state } i_x\right] \\
\doteq\ & P\left[(-\text{UCL} + (j_x - 1)g_2 - (1-r)c_{i_x})/r - \delta < X_{t1} - \delta < (-\text{UCL} + j_x g_2 - (1-r)c_{i_x})/r - \delta\right] \\
=\ & \Phi\left((-\text{UCL} + j_x g_2 - (1-r)c_{i_x})/r - \delta\right) - \Phi\left((-\text{UCL} + (j_x - 1)g_2 - (1-r)c_{i_x})/r - \delta\right),
\end{aligned}$$

where Φ is the cdf of standard normal distribution and $c_{i_x} = -\text{UCL} + (i_x - 0.5)g_2$.

Denote the transition matrix of Z_{t1} by $\mathbf{A}$. For the transition probability of $||\mathbf{Z}_{t2}||$, it can be easily set up by using Equations (1.9) and (1.10). Note that the degrees of freedom of the χ^2 distribution reduces to $p - 1$. Thus, the transition probability of $||\mathbf{Z}_{t2}||$ from transient state i_y to j_y is as follows. For $i_y, j_y = 0, 1, 2, \ldots, m_1$,

$$\begin{aligned}
& P\left[\ ||\mathbf{Z}_{t2}|| \text{ in state } j_y \mid ||\mathbf{Z}_{t2-1}|| \text{ in state } i_y\right] \\
=\ & \begin{cases} P\left\{(j_y - 0.5)^2 g_1^2/r^2 < \chi^2(p-1, c) < (j_y + 0.5)^2 g_1^2/r^2\right\} & \text{if } j_y \neq 0 \\ P\left\{\chi^2(p-1, c) < (0.5)^2 g_1^2/r^2\right\} & \text{if } j_y = 0. \end{cases}
\end{aligned} \tag{1.11}$$

Denote the transition matrix of $||\mathbf{Z}_{t2}||$ (Equation (1.11)) by $\mathbf{B}$. Note that as Figure (1.4) illustrates, some of the states have to be excluded from the set of transient states. Precisely, some of the transition probabilities are replaced by 0 and the states associated with those transition probabilities are regarded as absorbing states. Then, by using the Kronecker product ($\otimes$), we obtain $\mathbf{P}_1$ which is the transition matrix of the bivariate Markov chain

$\{Z_{t1}, ||\mathbf{Z}_{t2}||\}$.

$$\mathbf{P}_1 = \mathbf{A} \otimes \mathbf{B} \quad \text{under the condition } (\alpha - (m_2 + 1))^2 g_2^2 + \beta^2 g_1^2 < \text{UCL}^2,$$

where $\alpha = 1, 2, \ldots 2m_2 + 1$, and $\beta = 0, 1, \ldots m_1$ and the ordered pair (α, β) is a state of the Markov chain.

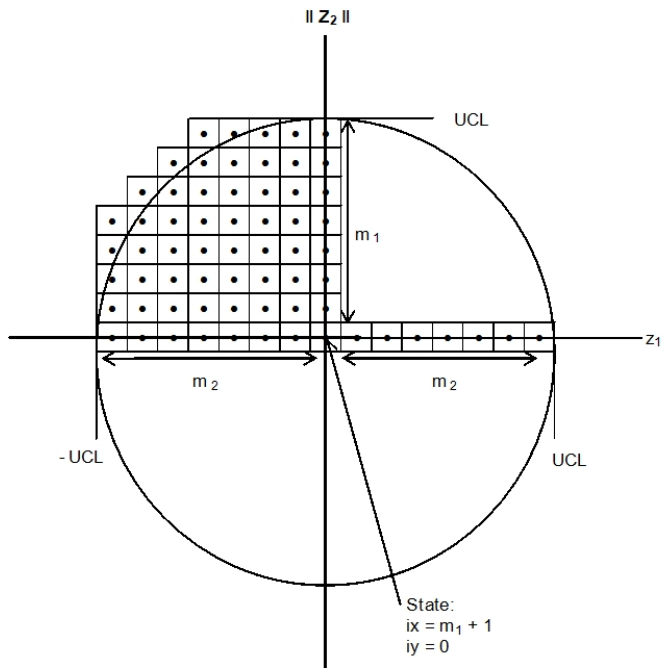


Figure 1.4: States in the Markov Chain Used for the Off-target Case of a MEWMA Chart (from Runger and Prabhu 1996).

1.6 Inertia effect

"Inertia" is the term used in classical physics, which was defined by Isaac Newton as his First Law of Motion in the *Principia*. Basically, when an object is at rest or in motion, it has a tendency to retain its state of rest or its velocity along a straight line unless it is acted upon by an external force. Similarly, "inertia" in statistical process control means

that when a mean shift occurs and the chart statistic is far away from the control limit, it takes a longer time for the chart to issue a signal. As a result, the chart statistic tends to stay within the control limits until large mean shifts are observed. In other words, "inertia" effect results in the potential delay in signaling. This problem happens to many control charts. It has been demonstrated by many researchers (Yashchin 1987, 1993; Ryan 2000; Lowry et al. 1992; Woodall and Maragah 1990; and Woodall and Mahmoud 2005) that the EWMA and MEWMA charts are heavily affected by this "inertia" effect when small smoothing parameters are employed because they do not weight the present values very heavily. Especially, to investigate the inertia effect, Woodall and Mahmoud (2005) proposed a new measure, so called the signal resistance, which is defined as the largest standardized deviation of the sample mean from the target value not leading to an immediate out-of-control signal.

1.7 Fast Initial Response (FIR) Feature

When a control chart is initiated due to start-up problems, a fast initial response (FIR) can be useful for rapid detection with a nonzero head start (HS) value. Lucas and Crosier (1982) and Lucas and Saccucci (1990) showed that this FIR feature is effective for the EWMA chart designed with a small smoothing parameter. When it comes to designing an EWMA chart, many researchers often use the asymptotic variance (Equation (1.3)) rather than the exact variance (Equation (1.2)) based on the assumption that the process will stay in control for a while and then shift to out-of-control. This might cause the control chart to respond slowly to an early shift. The reason is that when the control chart is designed with a small smoothing parameter, it takes a while for the exact variance (Equation (1.2)) to converge to the asymptotic variance (Equation (1.3)). As a result, the chart tends to be insensitive to start-up. Thus, in order to address this issue, an appropriate head start (HS) is suggested using the starting value (Z_0). Rhoads et al. (1996) suggested setting up

separate two one-sided EWMA charts with an appropriate HS. Clearly, this approach helps the charts to issue the out-of-control signal more quickly at start-up. However, one of the drawbacks of this approach for the EWMA chart is that two separate one-sided EWMA charts have to be implemented for monitoring. Moreover, it could be tricky to apply the FIR feature to multivariate EWMA charts.

1.8 Scope and Organization of the Thesis

This thesis provides various methods to improve the performance of the univariate EWMA and multivariate EWMA (MEWMA) charts.

When it comes to the smoothing parameters in the MEWMA control charts, the single smoothing parameter scheme supported by Lowry et al. (1992) is generally accepted. Its simplicity, combined with its high sensitivity to small and moderate process mean jumps, is at the core of its appeal. Lowry et al. (1992) advocated equal smoothing of each quality variable unless there is an a priori reason to weight quality characteristics differently. However, one may have situations where differential smoothing may be justified. In Chapter 2, we assess the performance of the differentially-smoothed MEWMA chart. The case of two quality variables (BEWMA) is discussed in detail. A bivariate Markov chain method that uses conditional distributions is developed for average run length (ARL) calculations. The proposed chart is shown to perform at least as well as Lowry et al. (1992)'s chart, and noticeably better in many mean jump directions. Comparisons with the recently introduced double-smoothed BEWMA chart and the use of univariate charts for the independent case show that the proposed differentially-smoothed BEWMA chart has superior performance.

As pointed out by several researchers (Stoumbos and Sullivan (2002) and Xie et al. (2011)), monitoring nonnormal data with the MEWMA chart often leads to a tricky situation where the out-of-control ARL becomes bigger than the in-control ARL. In Chapter 3, we use the Exponential-Exponential and the Gamma-Weibull joint distributions which

are constructed by the survival Gumbel copula (Gumbel 1960) function. We provide two monitoring methods and apply them to a univariate EWMA chart to enhance the chart performance as well as to stabilize the chart. The improvement over the existing method used by Xie et al. (2011) is discussed through extensive simulations. The Markov chain approach by the empirical CDF is introduced to validate the optimal ARL values obtained by simulations. How the charts respond to the change of correlation between bivariate random variables is also investigated in detail.

It is generally assumed that the process mean is fixed and known in evaluating chart performance. However, recently Ryu et al. (2010) introduced a new ARL-based performance measure, called expected weighted run length (EWRL) based on the assumption that the size of the mean shift is usually unknown and it follows a certain probability distribution. In Chapter 4, we generalize the new measure by adding another random variable T representing a transition point of the process mean shift. In addition, we suggest several new ARL-based performance measures to assess the performance of control charts more precisely when the process starts in-control initially and then a mean shift happens at some later time. We provide performance comparisons by applying our new ARL-based criteria to a variety of well-known univariate control charts.

Finally, Chapter 5 presents some conclusions and suggested directions for future research.

Chapter 2

Differential Smoothing in the BEWMA Chart

2.1 Introduction

The multivariate exponentially weighted moving average (MEWMA) control chart introduced by Lowry et al. (1992) is one of the most powerful control charts of its kind, offering high sensitivity to small and moderate process mean vector shifts. A measure of its impact is the 632 references to the paper recorded by Google Scholar (`http://scholar.google.ca/`, accessed on 22 May 2013). Calibrating the chart and assessing its performance was facilitated greatly by Runger and Prabhu (1996) and Prabhu and Runger (1997) who extended the Markov chain method to the MEWMA chart to approximate average run lengths.

We assume that the process data are $p \times 1$ random vectors $\mathbf{X}_1,\ \mathbf{X}_2,\ \mathbf{X}_3, ...$, each recording the p quality characteristics of interest, observed at sequential sampling periods. Further, we consider the case where the vectors are independent of one another and $\mathbf{X}_t$ is multivariate normally distributed with mean vector $\boldsymbol{\mu}_t = (\mu_{1t}, \mu_{2t}, ..., \mu_{pt})'$ and common covariance matrix $\boldsymbol{\Sigma}$, $t = 1, 2, 3, ...$ Without loss of generality, it is assumed that the in-control process mean is $\boldsymbol{\mu}_0 = (0, 0, ..., 0)' = \mathbf{0}$. Following Lowry et al. (1992), the MEWMA chart statistic

is

$$\mathbf{W}_t = r\mathbf{X}_t + (1-r)\mathbf{W}_{t-1}, \quad t = 1, 2, 3, ..., \tag{2.1}$$

where $\mathbf{W}_0 = \mathbf{0}$ and r is the smoothing parameter that determines the weights applied to present and past observation vectors ($0 < r \leq 1$). The chart signals when $\mathbf{W}_t'\mathbf{\Sigma}_{\mathbf{W}_t}^{-1}\mathbf{W}_t > H$ where H is the control limit and $\mathbf{\Sigma}_{\mathbf{W}_t}$ is the covariance matrix of $\mathbf{W}_t$.

In addition to its simplicity and clear generalization of the univariate EWMA chart, some advantages of this chart are: (a) average run length calculation is readily done using the Markov chain method, (b) average run lengths depend on $\boldsymbol{\mu}$ only through the non-centrality parameter $\delta = (\boldsymbol{\mu}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu})^{1/2}$, (c) optimal smoothing is possible, and (d) the chart has outstanding performance, faring well against its main competitors, the multivariate CUSUM and its variants (Woodall and Ncube 1985; Crosier 1988; Pignatiello and Runger 1990).

The focus of this article is whether gains in average run length are achieved by using varying degrees of smoothing on the quality variables. Thus, the chart statistic considered here is

$$\mathbf{W}_t = \mathbf{R}\mathbf{X}_t + (\mathbf{I}-\mathbf{R})\mathbf{W}_{t-1}, \quad t = 1, 2, 3, ..., \tag{2.2}$$

where $\mathbf{W}_0 = \mathbf{0}$, $\mathbf{R} = \text{diag}(r_1, r_2, ..., r_p)$, $0 < r_j \leq 1$, $j = 1, 2, ..., p$, and $\mathbf{I}$ is the $p \times p$ identity matrix. Naturally, some of the above listed features will be lost. However, there might be situations where varying amounts of smoothing may be justified. For instance: (a) departures in process mean may be different across quality variables, (b) some variables may evolve over time at a much different pace than other variables, and (c) the level of correlation between variables could vary substantially. One may argue that when two variables are highly correlated, the same degree of smoothing will suffice. It has also been suggested that when the variables are independent, separate univariate charts may be justified. It is of interest to notice that Lowry et al. (1992) were aware that in some situations differential weighting could be appropriate. In justifying their introduction and development of the the

equal smoothing case in Equation (2.1), they write, "If there is no a priori reason to weight past observations differently for the p variables being monitored, then $r_1 = r_2 = ... = r_p$."

Yumin (1996) and Hawkins et al. (2007) generalized the MEWMA control chart by modifying the smoothing matrix. Yumin (1996) advocated a standard MEWMA chart with a diagonal smoothing matrix applied after an orthogonal transformation of the data. The author suggested that this method should improve the standard MEWMA chart and recommended that further analysis be done to establish its numerical performance. Hawkins et al. (2007), on the other hand, discussed generalizing the standard MEWMA chart by employing a smoothing matrix with all diagonal elements equal and allowing non-zero but equal off-diagonal elements. Note that unlike Hawkins et al. (2007), our proposed differential smoothing scheme uses varying diagonal elements in the smoothing matrix and zeros for the off-diagonal entries.

In this thesis, we focus on the chart in Equation (2.2) when only two quality variables are measured ($p = 2$). Thus, we are dealing with the bivariate exponentially weighted moving average (BEWMA) chart with differential smoothing.

In Section 2.2, we describe the differentially-smoothed BEWMA chart in detail. In Section 2.3, Markov chain method for the differential smoothing is derived for the construction of the transition matrix. Extensive numerical performance comparisons with the standard single-smoothed BEWMA, double-smoothed BEWMA and the use of combined univariate EWMA charts are given in Section 2.4. The results of a sensitivity analysis to correlation misspecification are presented in Section 2.5. The chapter ends with a discussion of the advantages and limitations of the differentially-smoothed BEWMA chart in Section 2.6.

2.2 BEWMA Chart with Differential Smoothing

Suppose that two variables (X_1 and X_2), each representing a quality characteristic to be monitored, are observed from the process over time and let the 2×1 random vector $\mathbf{X}_t =$

$(X_{1t}, X_{2t})'$ represent the observations at time t. Further, we assume that $\mathbf{X}_t$ follows a bivariate normal distribution. When the process operates on-target,

$$\mathbf{X}_t \sim N_2\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_{\mathbf{X}}\right), \text{where } \mathbf{X}_t = \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix}, \boldsymbol{\mu}_0 = \begin{pmatrix} \mu_{10} \\ \mu_{20} \end{pmatrix}, \text{and } \boldsymbol{\Sigma}_{\mathbf{X}} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

When the process is out-of-control, we assume that the mean shifts to $\boldsymbol{\mu}_1 = (\mu_{11}, \mu_{21})'$ (or -$\boldsymbol{\mu}_1$) and the covariance matrix remains the same.

Taking $\boldsymbol{\mu}_0 = \mathbf{0} = (0,0)'$, the explicit form of the chart statistic from Equation (2.2) is

$$\mathbf{W}_t = \begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} + \begin{pmatrix} 1-r_1 & 0 \\ 0 & 1-r_2 \end{pmatrix} \mathbf{W}_{t-1}.$$

The chart signals when

$$Q_t = \mathbf{W}_t' \Sigma_{\mathbf{W}}^{-1} \mathbf{W}_t > H,$$

where $\mathbf{W}_0 = (0,0)'$ and $H(>0)$ is chosen to achieve a specified in-control ARL and $\boldsymbol{\Sigma}_{\mathbf{W}}$ is the asymptotic form of the covariance matrix of $\mathbf{W}_t$. Lowry et al. (1992) noted that the covariance matrix of $\mathbf{W}_t$ is

$$\boldsymbol{\Sigma}_{\mathbf{W_t}} = \begin{pmatrix} \dfrac{r_1\sigma_1^2[1-(1-r_1)^{2t}]}{2-r_1} & \dfrac{r_1 r_2 \rho\sigma_1\sigma_2[1-(1-r_1)^t(1-r_2)^t]}{r_1+r_2-r_1r_2} \\ \\ \dfrac{r_1 r_2 \rho\sigma_1\sigma_2[1-(1-r_1)^t(1-r_2)^t]}{r_1+r_2-r_1r_2} & \dfrac{r_2\sigma_2^2[1-(1-r_2)^{2t}]}{2-r_2} \end{pmatrix}.$$

Thus, the asymptotic covariance matrix $\boldsymbol{\Sigma}_{\mathbf{W}}$, obtained when $t \to \infty$, is

$$\boldsymbol{\Sigma}_{\mathbf{W}} = \begin{pmatrix} \dfrac{r_1\sigma_1^2}{2-r_1} & \dfrac{r_1 r_2 \rho\sigma_1\sigma_2}{r_1+r_2-r_1r_2} \\ \\ \dfrac{r_1 r_2 \rho\sigma_1\sigma_2}{r_1+r_2-r_1r_2} & \dfrac{r_2\sigma_2^2}{2-r_2}. \end{pmatrix}.$$

Moreover, the inverse asymptotic covariance matrix $\mathbf{\Sigma_W^{-1}}$ is

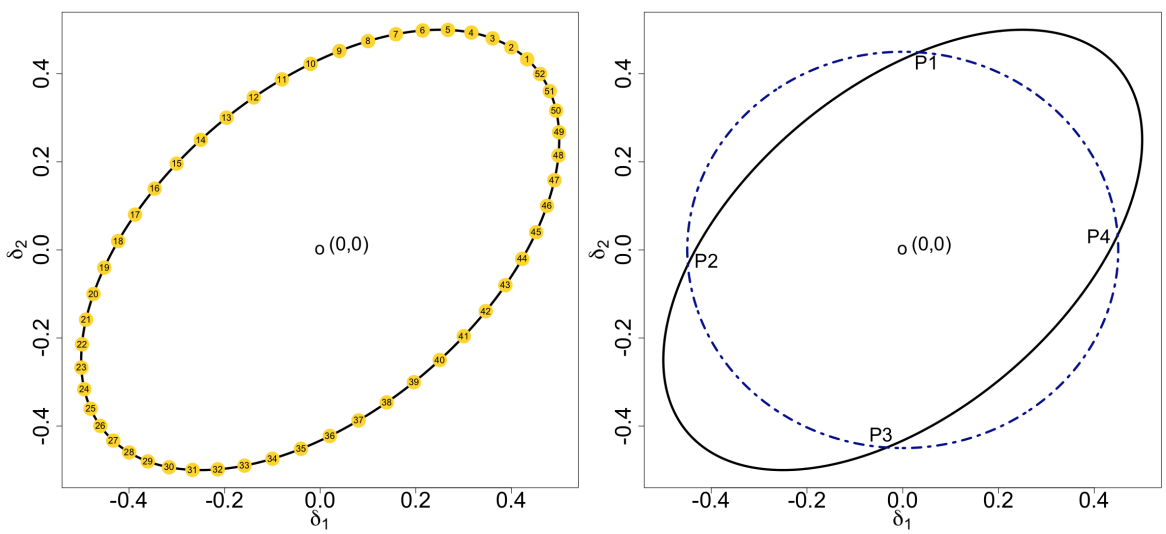
$$\mathbf{\Sigma_W^{-1}} = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

where

$$a = \frac{(2-r_1)(r_1+r_2-r_1r_2)^2}{r_1\sigma_1^2(r_1+r_2-r_1r_2)^2 - r_1^2r_2\rho^2\sigma_1^2(2-r_1)(2-r_2)},$$

$$b = \frac{\rho(2-r_1)(2-r_2)(r_1+r_2-r_1r_2)}{r_1r_2\rho^2\sigma_1\sigma_2(2-r_1)(2-r_2) - \sigma_1\sigma_2(r_1+r_2-r_1r_2)^2},$$

$$c = \frac{(2-r_2)(r_1+r_2-r_1r_2)^2}{r_2\sigma_2^2(r_1+r_2-r_1r_2)^2 - r_1r_2^2\rho^2\sigma_2^2(2-r_1)(2-r_2)}.$$



(a) (b)

Figure 2.1: (a) 52 $\boldsymbol{\mu}_1$ Points from the Ellipse Determined by the Non-centrality Parameter $\delta = \sqrt{\boldsymbol{\mu}_1'\mathbf{\Sigma_X}^{-1}\boldsymbol{\mu}_1}$. (b) The Intercepts of the Ellipse and Circle of Radius δ.

2.3 Transition Probability Matrix

The performance of a control chart is commonly measured by its ARL. Runger and Prabhu (1996) and Prabhu and Runger (1997) investigated the design of the standard MEWMA

control chart (see Equation (2.1)) through its ARL performance. They developed a procedure based on a bivariate Markov chain to approximate average run lengths. We denote their method by "Markov 1". In this section, we present a bivariate Markov chain method to approximate average run lengths for the proposed BEWMA chart in Equation (2.2). We denote this procedure by "Markov 2". The derivation of the new bivariate Markov chain method is provided. Note that in contrast with Markov 1, Markov 2 is based on conditional distributions.

Consider

$$\mathbf{W}_t'\Sigma_{\mathbf{W}}^{-1}\mathbf{W}_t = \begin{pmatrix} W_{1t} \\ W_{2t} \end{pmatrix}' \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} W_{1t} \\ W_{2t} \end{pmatrix} = aW_{1t}^2 + 2bW_{1t}W_{2t} + cW_{2t}^2 = H. \tag{2.3}$$

Equation (2.3) forms a tilted ellipse in W_{1t} and W_{2t} centered at (0,0). The ellipse's interior forms the in-control chart region. See Figure 2.2 to guide the process.

We present below the Markov chain representation of the BEWMA chart. We first determine the rectangle $[-\mathrm{UCL}_1, \mathrm{UCL}_1] \times [-\mathrm{UCL}_2, \mathrm{UCL}_2]$ that circumscribes the ellipse in order to obtain the Markov chain. We then divide $[-\mathrm{UCL}_1, \mathrm{UCL}_1]$ into $2m_1+1$ equal-length subintervals and similarly $[-\mathrm{UCL}_2, \mathrm{UCL}_2]$ into $2m_2+1$ equal-length subintervals. Note that if g_1 and g_2 denote the respective subinterval lengths, then $\mathrm{UCL}_i = (2m_i+1)g_i/2$, $i = 1, 2$. Differentiating Equation (2.3) with respect to W_{1t} and W_{2t} yields $g_1 = \frac{2\sqrt{Hc}}{(2m_1+1)\sqrt{ac-b^2}}$ and $g_2 = \frac{2\sqrt{Ha}}{(2m_2+1)\sqrt{ac-b^2}}$, respectively. Note that when $\mathbf{\Sigma_X} = \mathbf{I}$, and $r_1 = r_2 = r$, then g_1 and g_2 can be simplified to $\dfrac{2}{2m_1+1}\sqrt{\dfrac{r}{2-r}H}$ and $\dfrac{2}{2m_2+1}\sqrt{\dfrac{r}{2-r}H}$, respectively. This is the situation discussed by Runger and Prabhu (1996) in their Markov 1 procedure.

Next, consider the axis subintervals $\mathrm{I}_1, \ldots, \mathrm{I}_{2m_1+1}$ for W_{1t} and $\mathrm{J}_1, \ldots, \mathrm{J}_{2m_2+1}$ for W_{2t} from the above partitions. They induce $(2m_1+1) \times (2m_2+1)$ sub-rectangles $\mathrm{I}_i \times \mathrm{J}_j$, where $i = 1, \ldots, 2m_1+1$ and $j = 1, \ldots, 2m_2+1$. We denote the subinterval endpoints by $\mathrm{I}_i = [A_i, B_i]$ and $\mathrm{J}_j = [C_j, D_j]$. The center of $\mathrm{I}_i \times \mathrm{J}_j$ is $(\alpha_i, \beta_j) = (-\mathrm{UCL}_1 + (i - 0.5)g_1, -\mathrm{UCL}_2 + (j-0.5)g_2)$. The transient states of the proposed bivariate Markov chain

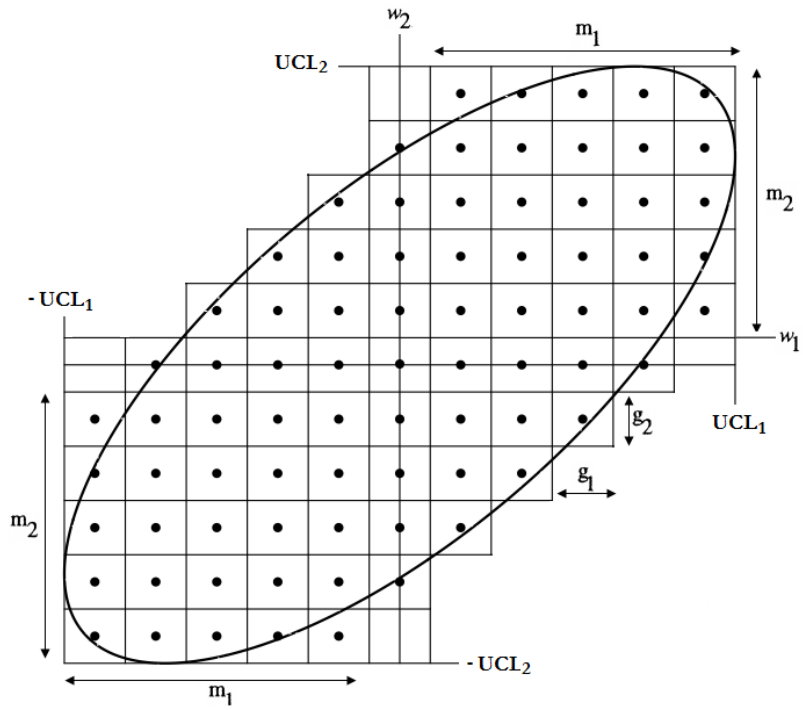


Figure 2.2: States in the Markov Chain Used for a Differentially-Smoothed BEWMA Chart. The Dots Represent the Centers of the States.

consist of all the subrectangles $\mathrm{I}_i \times \mathrm{J}_j$ whose centers (α_i, β_j) fall inside the ellipse, that is, inside the control region.

Consider transient states $\mathrm{I}_i \times \mathrm{J}_j$ and $\mathrm{I}_k \times \mathrm{J}_l$ where $i, k = 1, 2, \ldots, 2m_1 + 1$ and $j, l = 1, 2, \ldots, 2m_2 + 1$. It is shown that the transition probability (the off-target case) of $\mathbf{W}_t$ from state $\mathrm{I}_k \times \mathrm{J}_l$ to state $\mathrm{I}_i \times \mathrm{J}_j$ is given by

$$P\left(\mathbf{W}_t \in \mathrm{I}_i \times \mathrm{J}_j \middle| \mathbf{W}_{t-1} \in \mathrm{I}_k \times \mathrm{J}_l\right)$$

$$= P\left(\begin{pmatrix} W_{1,t} \\ W_{2,t} \end{pmatrix} \in \mathrm{I}_i \times \mathrm{J}_j \middle| \begin{pmatrix} W_{1,t-1} \\ W_{2,t-1} \end{pmatrix} \in \mathrm{I}_k \times \mathrm{J}_l\right)$$

$$= P\left(\begin{pmatrix} r_1 X_1 + (1 - r_1) W_{1,t-1} \\ r_2 X_2 + (1 - r_2) W_{2,t-1} \end{pmatrix} \in \mathrm{I}_i \times \mathrm{J}_j \middle| \begin{pmatrix} W_{1,t-1} \\ W_{2,t-1} \end{pmatrix} \in \mathrm{I}_k \times \mathrm{J}_l\right)$$

$$\doteq P\left(\begin{pmatrix} r_1X_1+(1-r_1)\alpha_k \\ r_2X_2+(1-r_2)\beta_l \end{pmatrix} \in \mathrm{I}_i \times \mathrm{J}_j\right),$$

where $\alpha_k = -\mathrm{UCL}_1 + (k-0.5)g_1$ and $\beta_l = -\mathrm{UCL}_2 + (l-0.5)g_2$.

$$P\left(\begin{pmatrix} r_1X_1+(1-r_1)\alpha_k \\ r_2X_2+(1-r_2)\beta_l \end{pmatrix} \in \mathrm{I}_i \times \mathrm{J}_j\right) = P\left(r_1X_1+(1-r_1)\alpha_k \in \mathrm{I}_i \;\&\; r_2X_2+(1-r_2)\beta_l \in \mathrm{J}_j\right)$$

$$= P\left(r_1X_1+(1-r_1)\alpha_k \in \mathrm{I}_i\right)\; P\left(r_2X_2+(1-r_2)\beta_l \in \mathrm{J}_j \middle| r_1X_1+(1-r_1)\alpha_k \in \mathrm{I}_i\right)$$

$$\doteq P\left(r_1X_1+(1-r_1)\alpha_k \in \mathrm{I}_i\right)\; P\left(r_2X_2+(1-r_2)\beta_l \in \mathrm{J}_j \middle| r_1X_1+(1-r_1)\alpha_k = \alpha_i\right)$$

$$= P\left(A_i < r_1X_1+(1-r_1)\alpha_k \le B_i\right)\; P\left(C_j < r_2X_2+(1-r_2)\beta_l \le D_j \middle| r_1X_1+(1-r_1)\alpha_k = \alpha_i\right)$$

$$= P\left(\frac{A_i-(1-r_1)\alpha_k}{r_1} < X_1 \le \frac{B_i-(1-r_1)\alpha_k}{r_1}\right)$$

$$\times P\left(\frac{C_j-(1-r_2)\beta_l}{r_2} < X_2 \le \frac{D_j-(1-r_2)\beta_l}{r_2}\middle| X_1 = \frac{\alpha_i-(1-r_1)\alpha_k}{r_1}\right).$$

When the process is on-target, we assume that $\mathbf{X}_t \sim N_2\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_{\mathbf{X}}\right)$, where $\boldsymbol{\mu}_0 = \begin{pmatrix} \mu_{10} \\ \mu_{20} \end{pmatrix} = \mathbf{0}$ and $\boldsymbol{\Sigma}_{\mathbf{X}} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. As a result, the conditional distribution of X_2 given X_1 is $X_2|X_1 = a \sim N\Big(\mu_{20} + \frac{\rho\sigma_2}{\sigma_1}(a-\mu_{10}), \sigma_2^2(1-\rho^2)\Big)$. To evaluate the off-target performance of the chart, let us assume that μ_{10} and μ_{20} shift to $\mu_{10}+\Delta x_1$ and $\mu_{20}+\Delta x_2$ respectively. Then the transition probability of $\mathbf{W}_t$ (off-target case) from state $\mathrm{I}_k \times \mathrm{J}_l$ to state $\mathrm{I}_i \times \mathrm{J}_j$ is

$$P\left(\mathbf{W}_t \in \mathrm{I}_i \times \mathrm{J}_j \middle| \mathbf{W}_{t-1} \in \mathrm{I}_k \times \mathrm{J}_l\right)$$

$$\doteq \left[\Phi\left(\frac{\frac{B_i-(1-r_1)\alpha_k}{r_1}-\mu_{10}-\Delta x_1}{\sigma_1}\right)-\Phi\left(\frac{\frac{A_i-(1-r_1)\alpha_k}{r_1}-\mu_{10}-\Delta x_1}{\sigma_1}\right)\right]$$

$$\times\left[\Phi\left(\frac{\frac{D_j-(1-r_2)\beta_l}{r_2}-\mu_{20}-\rho\frac{\sigma_2}{\sigma_1}\big(a-\mu_{10}\big)-\delta^*}{\sigma_2\sqrt{1-\rho^2}}\right)-\Phi\left(\frac{\frac{C_j-(1-r_2)\beta_l}{r_2}-\mu_{20}-\rho\frac{\sigma_2}{\sigma_1}\big(a-\mu_{10}\big)-\delta^*}{\sigma_2\sqrt{1-\rho^2}}\right)\right] \tag{2.4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $\delta^* = \Delta x_2 - \rho\frac{\sigma_2}{\sigma_1}\Delta x_1$ and $a = \dfrac{\alpha_i-(1-r_1)\alpha_k}{r_1}$. Since the transient states of the Markov chain are the sub-rectangles with centres (α,β) inside the ellipse in Figure 2.2, then

$$a\big(\alpha-(m_1+1)\big)^2 g_1^2 + 2b\big(\alpha-(m_1+1)\big)\big(\beta-(m_2+1)\big)g_1g_2 + c\big(\beta-(m_2+1)\big)^2 g_2^2 \le H, \tag{2.5}$$

where $\alpha = 1,2,\ldots,2m_1+1$ and $\beta = 1,2,\ldots,2m_2+1$. The transition probability matrix $\mathbf{P}$ consists of all the probabilities from Equations (2.4) and (2.5) for transient states. The off-taget average run length ARL_1 can be approximated by

$$ARL_1 \doteq \mathbf{s}'(\mathbf{I}-\mathbf{P})^{-1}\mathbf{1}, \tag{2.6}$$

where $\mathbf{s}$ is the starting column vector (with a 1 in the position of the transient state centered at (0,0), and 0 otherwise) and $\mathbf{1}$ is a column vector of 1s. A similar alternative to Equation (2.6) is derived in Appendix B.4 (see Equation (B.1)). Note that the dimensions of $\mathbf{P}$ and $\mathbf{s}$ depend on the number of transient states, and there are less than $(2m_1+1)\times(2m_2+1)$ such states.

Additionally, see details in Appendix B.1 for the derivation of transition probabilities in the case that $\boldsymbol{\mu}_0 = \begin{pmatrix}\mu_{10}\\ \mu_{20}\end{pmatrix} \neq \mathbf{0}$.

2.4 Performance Comparisons

In this section, we compare the performance of the standard BEWMA chart in Equation (2.1) with the proposed differentially-smoothed BEWMA chart in Equation (2.2) in terms of their optimal average run lengths. We also evaluate the performance of a recently developed variant of the standard MEWMA chart, the dMEWMA (Double Multivariate Exponentially Weighted Moving Average) chart. The basic idea is to smooth twice the chart statistic. Finally, for the case of independent quality indicators, we assess the performance of the chart obtained by combining separate univariate charts where signaling occurs as soon as one of the univariate charts signals.

Without loss of generality and for simplicity, here we assume that, when the process is on target, $\mathbf{X}_t \sim N_2\left(\boldsymbol{\mu}_0 = \mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{X}}\right)$, where $\boldsymbol{\Sigma}_{\mathbf{X}} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. This is equivalent to centering the original variables with respect to their on-target means and scaling them to have unit variances. Mean shifts will be denoted by $\boldsymbol{\mu}_1 = (\delta_1, \delta_2)'$, and by δ their distances from the origin in the non-centrality parameter scale, $\delta = \sqrt{\boldsymbol{\mu}_1' \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} \boldsymbol{\mu}_1}$.

To facilitate comparisons, chart performance will be evaluated on a selection of points $\boldsymbol{\mu}_1 = (\delta_1, \delta_2)'$ on the ellipse $\sqrt{\boldsymbol{\mu}_1' \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} \boldsymbol{\mu}_1} = \delta$ for several values of δ covering small to large process mean shifts. The coordinates δ_1 and δ_2 of $\boldsymbol{\mu}_1$ can be readily tracked through $\delta_1 = [\sqrt{1+|\rho|}\cos(\theta) + \sqrt{1-|\rho|}\sin(\theta)]\delta/\sqrt{2}$, $\delta_2 = [\sqrt{1+|\rho|}\cos(\theta) - \sqrt{1-|\rho|}\sin(\theta)]\delta/\sqrt{2}$, $0 \leq \theta \leq 2\pi$. Figure 2.1(a) shows 52 $\boldsymbol{\mu}_1$ points on the ellipse with $\rho = 0.50$ and $\delta = 0.50$. For the same ellipse, Figure 2.1(b) displays 4 reference points, labeled P1, P2, P3 and P4, corresponding to the intersection of the same ellipse with the circle $\sqrt{\boldsymbol{\mu}_1' \boldsymbol{\mu}_1} = \delta$. For the ρ and δ values considered, the reference points are P1 = (0.129, 0.483), P2= (−0.483, −0.129), P3 = −P1 and P4 = −P2. As will be seen later, these mean vector directions play a key role in the performance comparisons.

We follow the familiar approach and use the average run length as the basis to assess chart performance. Specifically, competing charts are first calibrated to have a specified

on-target average run length ARL_0. We have chosen $ARL_0 = 200$ for all the charts. Next, for a given mean vector shift, the off-target average run length ARL_1 is calculated for every chart. Finally, the chart with the smallest ARL_1 is selected as the best one for that mean shift. In the case of EWMA charts, one can additionally vary the smoothing parameters to search for the one that results in the smallest ARL_1 for the given departure, again, keeping ARL_0 at its set value. These charts are called **optimal charts** and their ARL values are denoted by ARL_{opt}. All comparisons here are based on ARL_{opt}.

2.4.1 Standard and Proposed BEWMA Charts

As described earlier, the standard BEWMA chart uses a single smoothing parameter r and its performance is determined by the non-centrality parameter $\delta = \sqrt{\boldsymbol{\mu}_1' \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} \boldsymbol{\mu}_{\mathbf{1}}}$. Thus, mean shifts $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are detected with the same average run length provided they produce identical non-centrality parameter values. In their Markov 1 average run length calculation, Runger and Prabhu (1996) first transform the quality indicators using $\boldsymbol{\Sigma}_{\mathbf{X}}^{-1/2}\mathbf{X}$, resulting in independent variables with zero means and unit variances. This transformation preserves the non-centrality parameter. Then they apply the smoothing (r) and evaluate chart performance on the transformed variables. We followed closely their well-described Markov 1 approach for calculating ARLs, obtaining optimal ARL values for every mean departure δ by varying the smoothing parameter r.

Similarly, for the proposed BEWMA chart with differential smoothing, ARL calculations were carried out using Markov 2 (see Section 2.3). For a given correlation coefficient ρ and mean shift $\boldsymbol{\mu}_1 = (\delta_1, \delta_2)'$, the smoothing parameters r_1 and r_2 were varied to identify the optimal off-target ARL_{opt}. It is interesting to note that, although chart performance is not the same for a given non-centrality parameter distance, there are some obvious simplifications. Specifically, the optimal performance on mean shift (δ_1, δ_2) is identical to that for

$(-\delta_1, -\delta_2)$ for same value of ρ.

Note that both Markov 1 and Markov 2 involve bivariate Markov chains with the number of transient states determined by the pair of integers (m_1, m_2). After carefully examining the effect of different choices for m_1 and m_2, we decided to use $m_1 = m_2 = 30$ for both methods (see Appendix B.3). This choice provided stable ARL values throughout. The number of transient states varies from 1,700 to 3,000 depending on the value of ρ.

2.4.2 Double Bivariate Exponentially Weighted Moving Average (dBEWMA) Chart

First proposed in its univariate version by Shamma et al. (1991) and Shamma and Shamma (1992), the double EWMA chart was also studied by Zhang and Chen (2005) and Mahmoud and Woodall (2010), and then extended to its multivariate form by Alkahtani and Schaffer (2012). We focus on the bivariate case, (dBEWMA). The basic idea is to smooth twice the chart statistic. Specifically, the chart statistic $\mathbf{W}_t$ is given by

$$\left.\begin{aligned} \mathbf{W}_t &= r\mathbf{Z}_t + (1-r)\mathbf{W}_{t-1}, \\ \mathbf{Z}_t &= r\mathbf{X}_t + (1-r)\mathbf{Z}_{t-1}, \end{aligned}\right\} \quad t = 1, 2, ...$$

Alkahtani and Schaffer (2012) considered initially the possibility of using different smoothing parameters, but they only addressed the case of identical smoothing for both stages. Regarding the ARL calculation, there is no Markov chain method available, and so the authors relied on simulations. In Alkahtani and Schaffer (2012), the authors compared the performance of their double smoothing chart with the standard MEWMA chart for selected mean shifts and r values. Similar to the standard MEWMA chart, the chart depends on a mean shift $\boldsymbol{\mu}_1 = (\delta_1, \delta_2)'$ only through the non-centrality parameter $\delta = \sqrt{\boldsymbol{\mu}_1' \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} \boldsymbol{\mu}_1}$.

Here we focus on the optimal ARL obtained by searching for the smoothing amount r

that yields the smallest off-target ARL for given ρ and δ. Our ARL_{opt} calculations were each based on 300,000 simulated charts.

2.4.3 Multiple Univariate Charts for Independent Quality Indicators

When the quality indicators are independent, the benefit of a single multivariate chart may be diminished since borrowed strength arising from cross correlations is not available. It has been suggested that using individual univariate charts may be more efficient. In this case, the monitoring plan issues an out-of-control signal as soon as one of the univariate charts does so. A bonus of the univariate charts is that diagnosing which quality indicator triggers the out-of-control process excursion is readily available, a constant challenge for multivariate charts. See Montgomery (2013, Ch. 10) for a good discussion on the use of multiple univariate charts.

For two independent quality indicators, X_{1t} and X_{2t}, we apply first the Markov chain method to each variable (see, e.g. Lucas and Saccucci 1990), denote by RL_1 and RL_2 the associated run lengths until signal, and by $\mathbf{P}_1$ and $\mathbf{P}_2$ the respective transition matrices. Then, $P(RL_i > n) = \mathbf{s}_i'\mathbf{P}_i^n\mathbf{1}_i$, $i = 1, 2$. By the product rule under independence, for the run length RL of the combined chart, the probability of signaling after n runs is

$$P(RL > n) = P(\min\{RL_1, RL_2\} > n) = \prod_{i=1}^{2} \mathbf{s}_i'\mathbf{P}_i^n\mathbf{1}_i, \quad n = 0, 1, 2, ...$$

Hence, the average run length for the combined chart is

$$ARL_1 = \sum_{n=0}^{\infty} \prod_{i=1}^{2} \mathbf{s}_i'\mathbf{P}_i^n\mathbf{1}_i. \tag{2.7}$$

Varying r_1 and r_2 to minimize ARL_1 produces the optimal average run length (ARL_{opt}) for the combined procedure.

The values of ARL_{opt} were obtained using 2,000 terms in the sum of Equation (2.7) for all the mean shifts and values of ρ considered. Note that this chart depends on the specific δ_1 and δ_2 in the mean shift, and thus it is not directionally invariant.

2.4.4 Numerical Results

1. Table 2.1 contains an example of the optimal average run lengths (ARL_{opt}) obtained using the proposed differentially-smoothed BEWMA chart. The chosen mean shifts $\boldsymbol{\mu}_1 = (\delta_1, \delta_2)'$ are the 52 points from Figure 2.1(a) on ellipse $\sqrt{\boldsymbol{\mu}_1' \boldsymbol{\Sigma}_{\mathbf{X}}^{-1} \boldsymbol{\mu}_1} = \delta = 0.50$ with correlation $\rho = 0.50$. The reference points P1, P2, P3 and P4 are noted. The table reflects the symmetries in ARL_{opt} behavior noted earlier. Clearly, there is variation in ARL_{opt}, which ranges from 24.68 to 26.69. The optimal average run length for the standard BEWMA chart obtained from Markov 1 is 26.68. In this particular case, the percentage reduction in ARL with the differential smoothing method is between -0.04% to 7.50%. Thus, apart from differences due to numerical accuracy, none of the ARL_{opt} values exceeds the optimal average run length for the standard BEWMA chart which the proposed chart achieves when the quality indicators experience the same individual jump in mean, i.e., when $\delta_1 = \delta_2$. More strikingly, the proposed method performs better in many directions, reaching its maximum gain in efficiency (around 7.50%) in directions near the reference points.

2. Tables similar to Table 2.1 were produced for other values of ρ and δ. Figures 2.3 and 2.4 compare 52 optimized differential BEWMA charts to the one standard BEWMA chart for a variety of ρ and δ values. Also plotted are the optimal ARLs for the standard BEWMA and for the double-smoothed dBEWMA charts obtained, respectively, using Markov 1 and simulation, as described earlier. Since the latter charts are directionally invariant, their ARL values are the same for all the 52 directions

Table 2.1: Mean Shifts (δ_1, δ_2) and ARL_{opt} Values of 52 Extracted Points ($\rho = 0.50, \delta = 0.50$)

Point	(δ_1, δ_2)	ARL_{opt}	Point	(δ_1, δ_2)	ARL_{opt}
# 1	(0.433, 0.433)	26.69	# 27	(-0.433, -0.433)	26.69
# 2	(0.400, 0.460)	26.47	# 28	(-0.400, -0.460)	26.47
# 3	(0.361, 0.480)	25.68	# 29	(-0.361, -0.480)	25.68
# 4	(0.316, 0.494)	25.09	# 30	(-0.316, -0.494)	25.09
# 5	(0.267, 0.500)	24.78	# 31	(-0.267, -0.500)	24.78
# 6	(0.236, 0.499)	24.74	# 32	(-0.236, -0.499)	24.74
# 7	(0.158, 0.490)	24.95	# 33	(-0.158, -0.490)	24.95
P1	(0.129, 0.483)	25.15	P3	(-0.129, -0.483)	25.15
# 8	(0.100, 0.474)	25.42	# 34	(-0.100, -0.474)	25.42
# 9	(0.040, 0.452)	26.08	# 35	(-0.040, -0.452)	26.08
# 10	(-0.020, 0.423)	26.35	# 36	(0.020, -0.423)	26.35
# 11	(-0.080, 0.387)	26.54	# 37	(0.080, -0.387)	26.54
# 12	(-0.139, 0.346)	26.65	# 38	(0.139, -0.346)	26.65
# 13	(-0.196, 0.300)	26.70	# 39	(0.196, -0.300)	26.70
# 14	(-0.250, 0.250)	26.70	# 40	(0.250, -0.250)	26.70
# 15	(-0.300, 0.196)	26.70	# 41	(0.300, -0.196)	26.70
# 16	(-0.346, 0.139)	26.65	# 42	(0.346, -0.139)	26.65
# 17	(-0.387, 0.080)	26.54	# 43	(0.387, -0.080)	26.54
# 18	(-0.423, 0.020)	26.35	# 44	(0.423, -0.020)	26.35
# 19	(-0.452, -0.040)	26.08	# 45	(0.452, 0.040)	26.08
# 20	(-0.474, -0.100)	25.39	# 46	(0.474, 0.100)	25.39
P2	(-0.483, -0.129)	25.11	P4	(0.483, 0.129)	25.11
# 21	(-0.490, -0.158)	24.91	# 47	(0.490, 0.158)	24.91
# 22	(-0.499, -0.236)	24.68	# 48	(0.499, 0.236)	24.68
# 23	(-0.500, -0.267)	24.73	# 49	(0.500, 0.267)	24.73
# 24	(-0.494, -0.316)	25.04	# 50	(0.494, 0.316)	25.04
# 25	(-0.480, -0.361)	25.61	# 51	(0.480, 0.361)	25.61
# 26	(-0.460, -0.400)	26.46	# 52	(0.460, 0.400)	26.46

and thus show up in the plots as horizontal lines. Comparing the performance of the proposed differentially-smoothed BEWMA chart with that of the standard BEWMA chart, similar patterns to those observed from Table 2.1 emerge. Namely, the differentially-smoothed BEWMA chart is as good as the standard BEWMA chart, and better in many directions. Regarding the double-smoothed dBEWMA chart, for small mean shifts it performs worse than the standard BEWMA chart (and hence than the proposed differentially-smoothed chart) and a bit better for large mean jumps. In the latter case, only in a few directions the chart does better than the proposed chart. The noted patterns in performance of the dBEWMA chart relative to the stan-

dard BEWMA chart is consistent with the findings in the univariate case reported by Mahmoud and Woodall (2010).

3. We focus now on the actual smoothing parameters for the proposed differentially-smoothed BEWMA chart that yield the optimal ARLs. We denote the pair by $\mathbf{r}_{opt} = (r_1, r_2)'$. Figure 2.5 displays the pairs $\mathbf{r}_{opt}$ for the 52 directions in the ellipse from Figure 2.1(a) for representative correlation levels ρ and mean jumps specified by δ. The line $r_1 = r_2$, which matches the equal smoothing case corresponding to the standard BEWMA chart, is also drawn. Overall, for each correlation level ρ, similar patterns in $\mathbf{r}_{opt}$ emerge as δ moves from small to large values, however, the variation in $\mathbf{r}_{opt}$ increases. The salient point is that r_1 and r_2 differ the most when the mean directions are near the reference points P1, P2, P3 and P4, in this case the proposed differentially-smoothed BEWMA chart outperforms the standard BEWMA chart. On the other hand, $r_1 \approx r_2$ for mean directions where $\delta_1 \approx \delta_2$, and in this case, the performance of the proposed chart is similar to the standard BEWMA chart.

4. Finally, consider the case of independent quality indicators X_1 and X_2, i.e., $\rho = 0$. In addition to the performances of the three charts so far examined, we also considered the combined univariate EWMA charts whose average run length is calculated from Equation (2.7). To facilitate comparisons, the optimal average run length for the combined chart was calculated for the 52 mean shifts from the ellipse displayed in Figure 2.1(a) for two values of δ, namely $\delta = 0.25,\ 2.0$. This chart is direction-dependent. The resulting optimal ARLs are plotted in Figure 2.6 along with those for the other three charts examined here. Compared to the standard BEWMA chart, the combined chart performs better in some directions and much worse in others. When δ is large, the standard BEWMA chart is superior in most directions. As for the differentially-smoothed BEWMA chart, the proposed chart shows superior performance in all directions for mean vector jumps of any size.

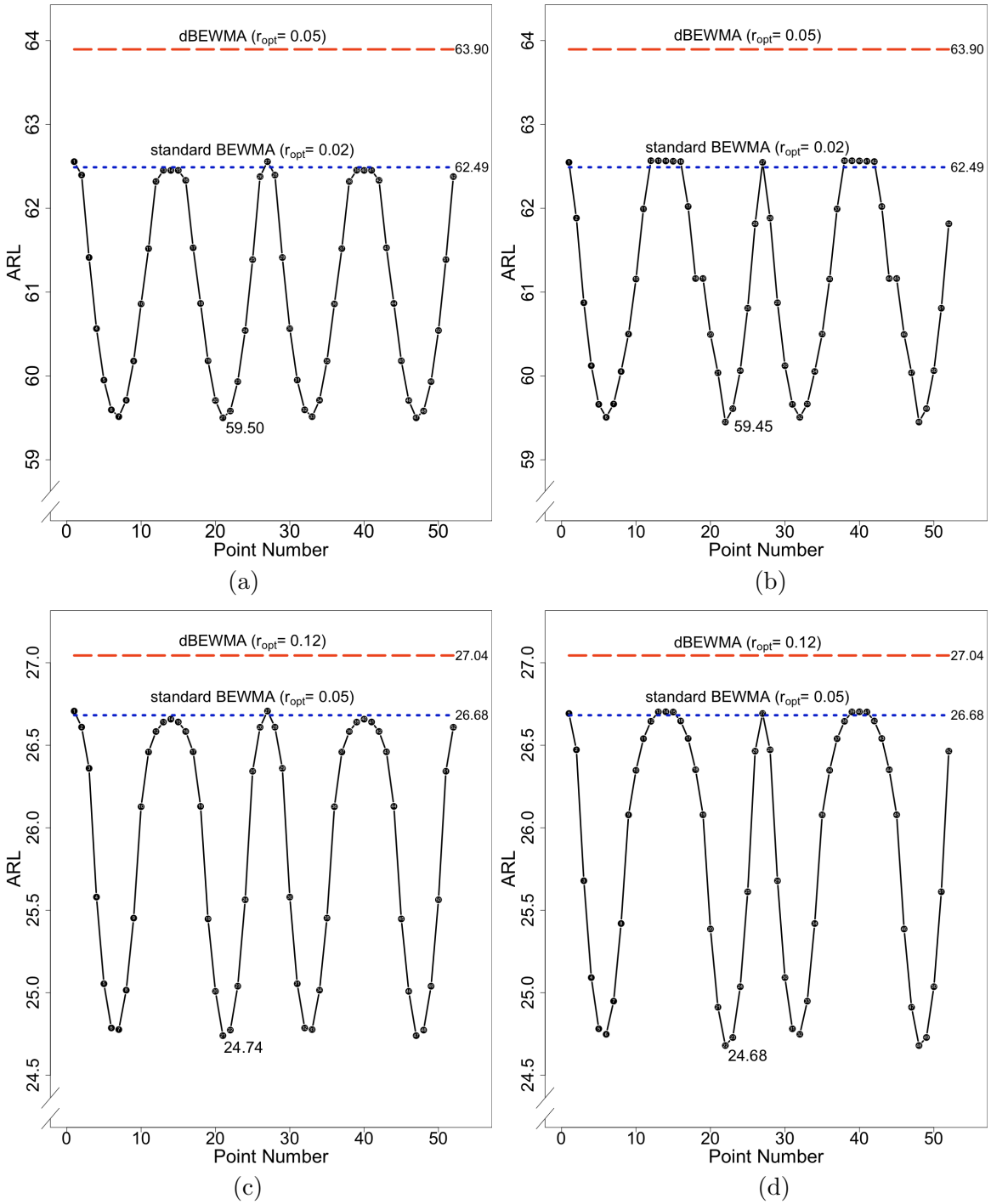


Figure 2.3: Optimal Average Run Lengths (ARL_{opt}) for the 52 Mean Jumps on the Ellipse (a) $\rho = 0.25$, $\delta = 0.25$; (b) $\rho = 0.50$, $\delta = 0.25$; (c) $\rho = 0.25$, $\delta = 0.50$; and (d) $\rho = 0.50$, $\delta = 0.50$.

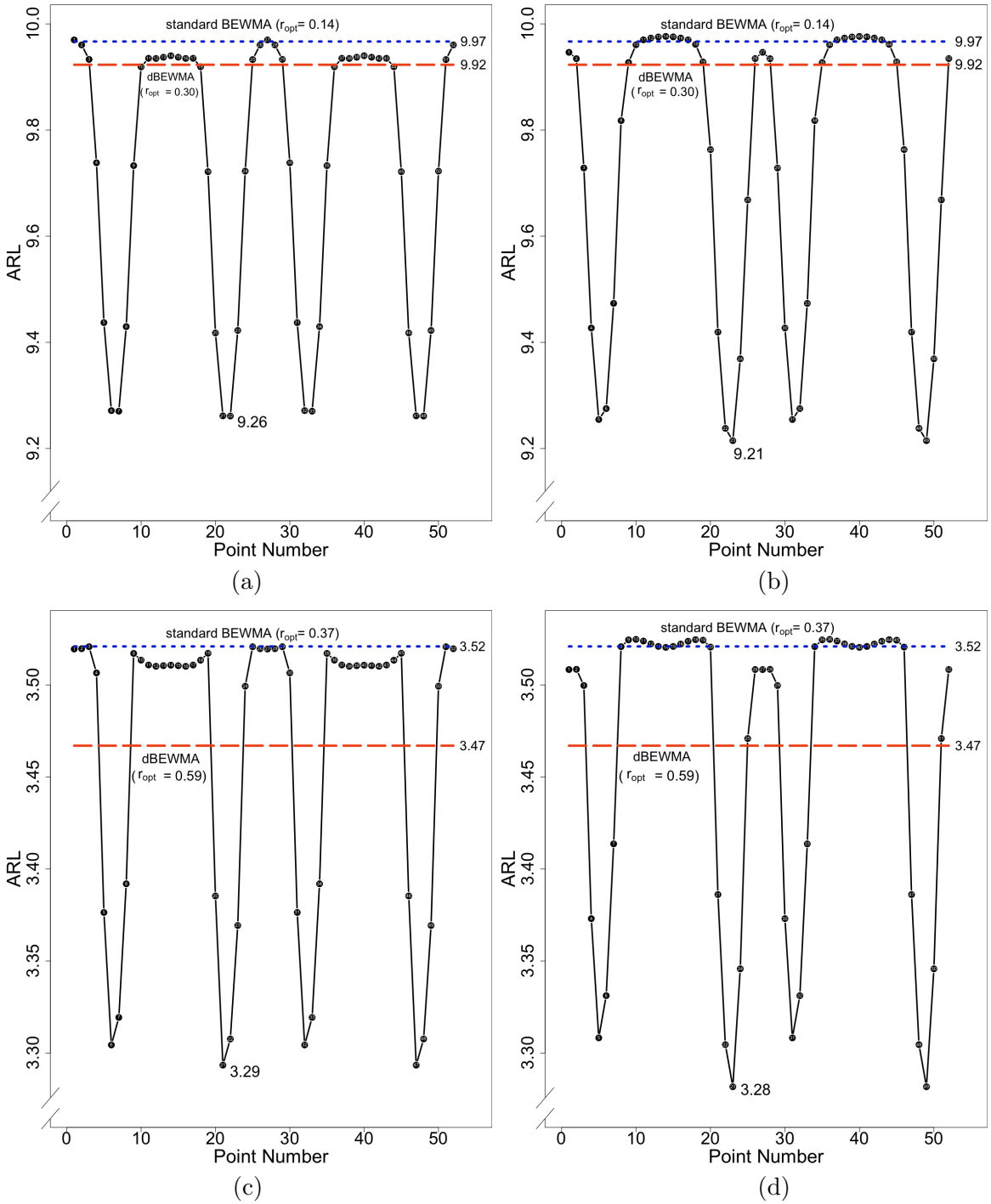


Figure 2.4: Optimal Average Run Lengths (ARL_{opt}) for the 52 Mean Jumps on the Ellipse (a) $\rho = 0.25$, $\delta = 1.00$; (b) $\rho = 0.50$, $\delta = 1.00$; (c) $\rho = 0.25$, $\delta = 2.00$; and (d) $\rho = 0.50$, $\delta = 2.00$.

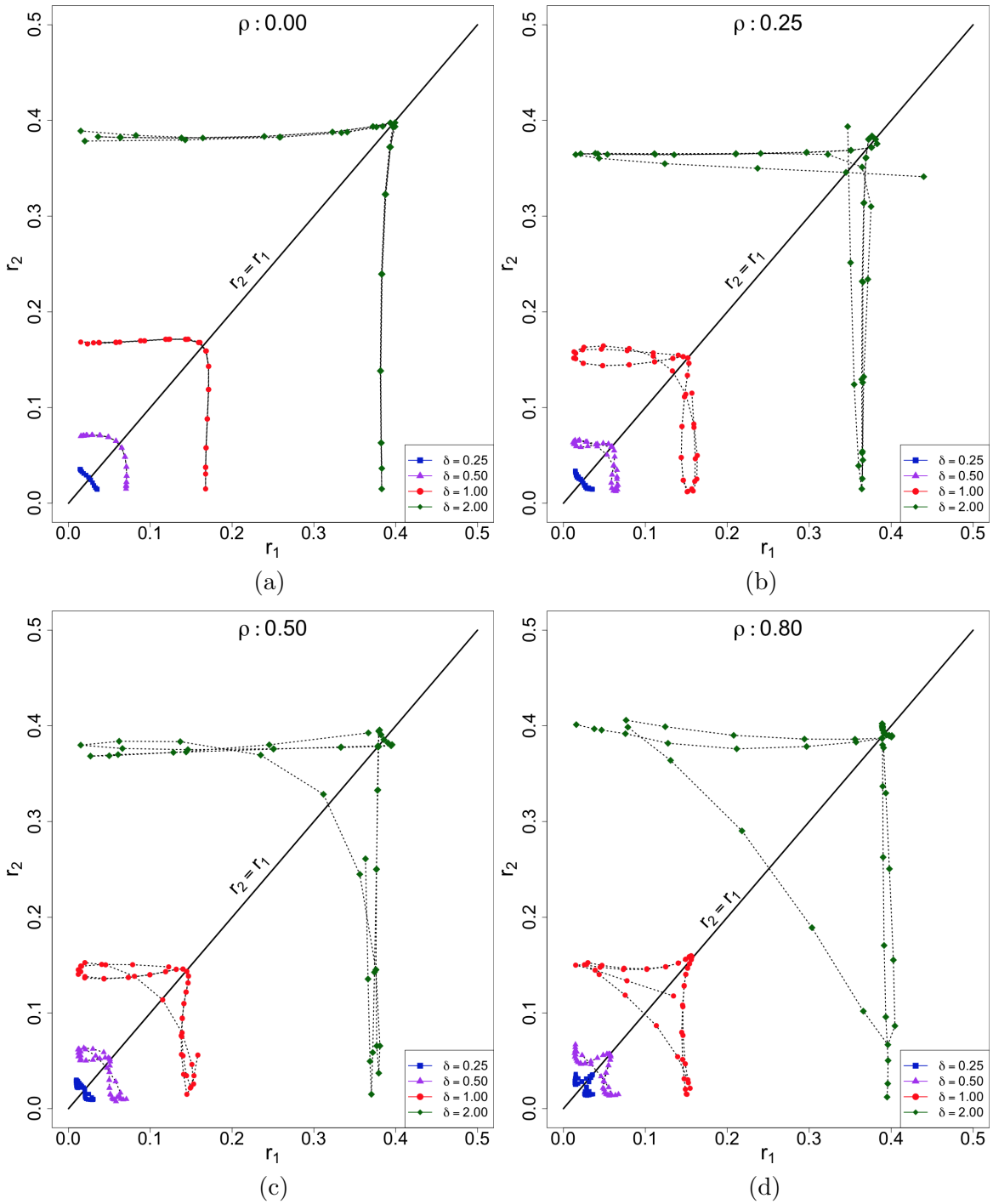


Figure 2.5: Optimal Smoothing Amounts r_1 and r_2 for the Differentially-Smoothed Chart for (a) $\rho = 0.0$, (b) $\rho = 0.25$, (c) $\rho = 0.50$, and (d) $\rho = 0.80$.

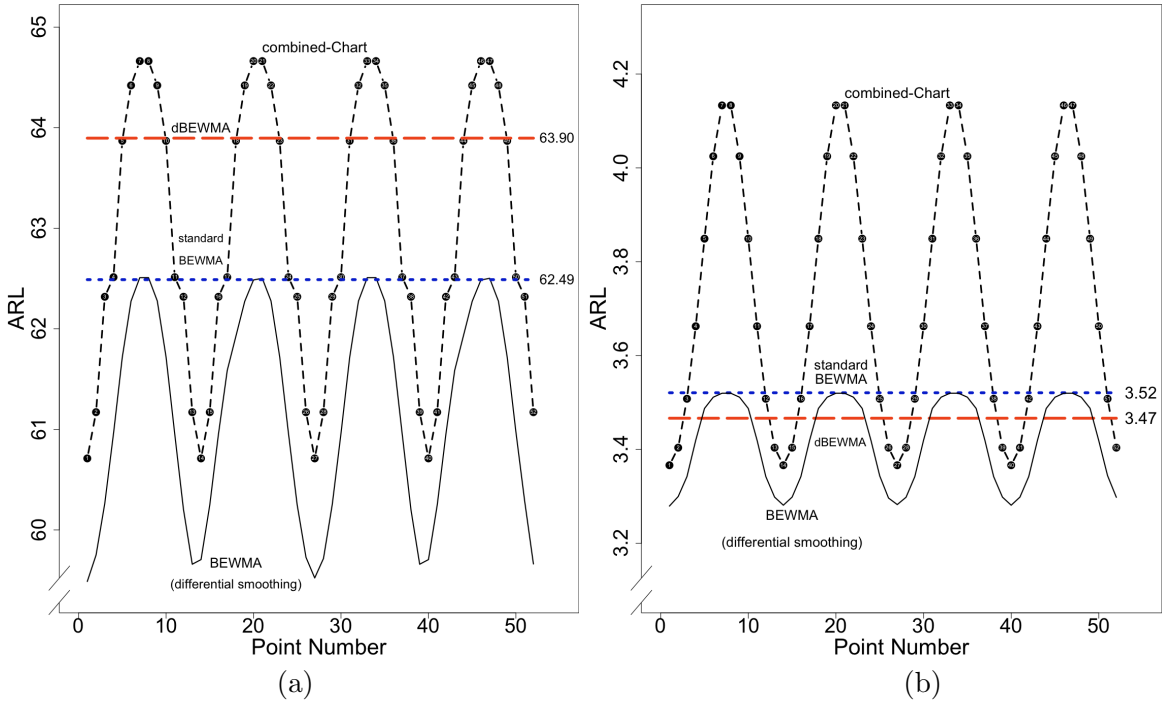


Figure 2.6: Optimal Average Run Length Performance for all the Four Charts Discussed for the Case of Independent Quality Indicators for (a) $\delta = 0.25$ and (b) $\delta = 2.0$.

2.5 Sensitivity Analysis

One issue of interest is the robustness of the differentially-smoothed BEWMA chart to misspecification of the correlation coefficient ρ. Specifically, suppose the chart calibration is done using the value of ρ for the correlation coefficient, but the true correlation is $\rho + \epsilon$ or $\rho - \epsilon$. What is the effect of the misspecification on the ARL? Recall that ρ is treated as a process parameter requiring estimation from Phase I process data.

Tables 2.2 and 2.3 present the results of a sensitivity study using six correlation misspecification cases: $0.20 \pm 0.05, 0.50 \pm 0.05$ and 0.8 ± 0.05. For instance, the true correlation may be $0.20 + 0.05 = 0.25$ but the chart is calibrated using 0.20. The Markov chain method was used to compute the average run lengths based on the misspecified and true correlation coefficients using the control limit from the misspecified correlation. To assess the effect of the misspecification, we used the absolute relative deviation percentage (ARDP) designed

by

$$\text{ARDP} = 100\frac{|ARL^{(M)} - ARL^{(T)}|}{ARL^{(M)}}\%,$$

where $ARL^{(M)}$ and $ARL^{(T)}$ are the average run lengths from the misspecified and true correlations, respectively.

For the in-control situation (see Table 2.2), the effect has little impact on average run length performance. However, as Table 2.3 reveals, the difference in average run length increases with increasing correlations for out-of-control scenarios. Note that the smaller the jump in the mean, the smaller the difference in average run length. At the same time, the direction also plays an important role.

Table 2.2: On-target ARLs for Various Misspecified Correlations

$ARL_0 = 200$, $(r_1, r_2) = (0.3, 0.2)$						
Correlation	Low ($\rho = 0.20 \pm 0.05$)		Moderate ($\rho = 0.50 \pm 0.05$)		High ($\rho = 0.80 \pm 0.05$)	
ρ	$\rho = 0.15$	$\rho = 0.25$	$\rho = 0.45$	$\rho = 0.55$	$\rho = 0.75$	$\rho = 0.85$
ARL_0	198.39	201.29	197.65	201.28	196.58	198.42
ARDP	0.81%	0.65%	1.18%	0.64%	1.71%	0.79%

2.6 Discussion

It has been 20 years since the standard equal smoothing MEWMA chart was introduced. Its simplicity, combined with its good sensitivity to mean vector process jumps, has made the chart a favorite in applied work. Montgomery (2009), arguably the best and most popular book on quality control methods, places it as the leading multivariate chart.

It is important to point out that Lowry et al. (1992) were aware that in some situations differential weighting could be appropriate. In justifying their introduction and development of the equal smoothing case, they write, "If there is no a priori reason to weight past observations differently for the p variables being monitored, then $r_1 = r_2 = ... = r_p$."

The question is: under what practical circumstances one would prefer differential smoothing over equal smoothing? Adopting the optimal average run length as a criterion, the results

Table 2.3: Off-target ARLs for Various Misspecified Correlations and Several Mean Jumps

$ARL_0 = 200$, $(r_1, r_2) = (0.3, 0.2)$					
Departure	Low ($\rho = 0.20 \pm 0.05$)				
(δ_1, δ_2)	$\rho = 0.20$	$\rho = 0.15$	ARDP	$\rho = 0.25$	ARDP
(0.0, 0.4)	50.27	50.76	0.97%	49.58	1.37%
(0.0, 0.2)	118.89	118.97	0.07%	118.43	0.39%
(0.2, 0.0)	131.58	131.55	0.02%	131.32	0.20%
(0.2, 0.2)	100.78	97.68	3.08%	103.15	2.35%
(0.2, 0.4)	49.47	48.13	2.71%	50.41	1.90%
(0.4, 0.0)	60.63	61.21	0.96%	59.90	1.20%
(0.4, 0.2)	56.60	54.74	3.29%	58.02	2.51%
(0.4, 0.4)	36.77	35.10	4.54%	38.09	3.59%
Departure	Moderate ($\rho = 0.50 \pm 0.05$)				
(δ_1, δ_2)	$\rho = 0.50$	$\rho = 0.45$	ARDP	$\rho = 0.55$	ARDP
(0.0, 0.4)	41.30	43.10	4.36%	39.24	4.99%
(0.0, 0.2)	106.79	108.78	1.86%	104.29	2.34%
(0.2, 0.0)	120.21	121.71	1.25%	117.58	2.19%
(0.2, 0.2)	112.10	109.48	2.34%	113.85	1.56%
(0.2, 0.4)	51.95	51.71	0.46%	51.66	0.56%
(0.4, 0.0)	50.22	52.15	3.84%	47.69	5.04%
(0.4, 0.2)	62.46	61.40	1.70%	62.76	0.48%
(0.4, 0.4)	44.44	43.02	3.20%	45.56	2.52%
Departure	High ($\rho = 0.80 \pm 0.05$)				
(δ_1, δ_2)	$\rho = 0.80$	$\rho = 0.75$	ARDP	$\rho = 0.85$	ARDP
(0.0, 0.4)	22.78	26.54	16.51%	18.23	19.97%
(0.0, 0.2)	72.59	80.40	10.76%	60.77	16.28%
(0.2, 0.0)	85.28	93.17	9.25%	72.77	14.67%
(0.2, 0.2)	120.10	117.17	2.44%	120.89	0.66%
(0.2, 0.4)	42.33	45.15	6.66%	37.48	11.46%
(0.4, 0.0)	27.25	31.89	17.03%	21.61	20.70%
(0.4, 0.2)	57.32	58.99	2.91%	53.38	6.87%
(0.4, 0.4)	50.85	49.32	3.01%	51.85	1.97%

reported in this chapter demonstrate that noticeable gains in optimal average run length are achieved using differential smoothing when there is a substantial difference in size among the individual variable mean jumps one wishes to detect.

For example, consider the case of two independent quality variables X_1 and X_2, and specific mean jump $\delta = \sqrt{\boldsymbol{\mu}_1' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1} = 2.0$, where $\boldsymbol{\mu}_1' = (\delta_1, \delta_2) = (0.20, 1.99)$. The optimal weight for a univariate EWMA chart for X_1 to detect $\delta_1 = 0.20$ with $ARL_0 = 200$ is $r_1 = 0.02$. Similarly, the optimal weight to be used on X_2 to detect $\delta_2 = 1.99$ in a univariate EWMA chart is $r_2 = 0.43$. This discrepancy between r_1 and r_2 suggests that it will be unlikely that when the process is monitored by using a BEWMA chart for (X_1, X_2), the optimal weights to be used will be identical.

One issue of interest is how much detection power is lost when the proposed chart is calibrated for a certain mean vector departure (δ_1, δ_2), but the mean vector departure that occurs is (δ_1', δ_2'). Let us consider one extreme case where the differentially-smoothed BEWMA chart is designed to detect a mean vector shift for which one of the components does not change at all. Under this setting, the smoothing parameter for the variable with the unchanging mean will be very small. In this case, the correlation between the two variables plays an important role in the chart performance when the chart is used in other directions. When there is no correlation, the proposed chart performs poorly in every direction. However when the correlation increases, the chart tends to perform better in many directions. More importantly, for large correlations, the proposed chart expands the area where it is competitive.

The Markov chain method developed here for average run length calculation is accurate and efficient for the case $p = 2$ (BEWMA chart). The conditional argument to calculate transition probabilities extends to any number of quality variables p. However, the number of transient states grows quickly with p. If p is large, simulations may be preferred for computing ARLs.

A numerical challenge posed by the proposed method is finding the optimal values for

the weighting. The standard chart uses only one level of smoothing, and so one needs to optimize over one variable (r) only. However, when differential smoothing is applied, optimization is required on several variables, namely r_1, r_2,...,r_p. It just takes a bit longer to obtain numerically the optimal values.

Although not explored here, Rigdon (1995a, b) provided alternative analytical methods for obtaining the ARL of the standard MEWMA control chart by solving single and double integral equations. Note that the numerical methods are based on equal smoothing parameters and they are directionally invariant, but the differential smoothing scheme is not. It will be interesting to investigate the adaptation required to make the methods work under the differentially-smoothed scheme.

Finally, an important problem not addressed in this chapter is that of chart inertia. The problem was noted first and partially addressed by Lowry et al. (1992) in the context of the standard MEWMA chart. Woodall and Mahmoud (2005) presented a careful study on chart inertia, covering many of the charts available. The problem arises when, after a period of normal operation, the process goes off-target at a sampling period where the last value of the chart statistic is far away from the chart control limit. If the chosen smoothing parameter is small, the chart can take an unduly long time to reach the control limit thus becoming inefficient at detecting the change in the process mean. In this situation, the use of Hotelling's χ^2 limits in conjunction with MEWMA charts may be more effective. However, preliminary findings reveal that one can increase the smoothing amounts r_1 and r_2 a bit without causing a substantial increase in the off-target average run length. This provides a limited way to address adverse inertia effects. But we agree with the general recommendation of Lowry et al. (1992) and Woodall and Mahmoud (2005) that supplementing the differentially-smoothed chart with a Hotelling's χ^2 chart will increase its efficiency. Naturally, some ARL numerical work would be needed to tune the two charts to calibrate the combined chart to have a desired on-target ARL. This issue will certainly be worth considering further.

Chapter 3

EWMA Chart for Monitoring Bivariate Skewed Data

3.1 Introduction

When monitoring multivariate process data, many well-known multivariate control charts (Hotelling χ^2 charts (Hotelling 1947), the multivariate exponential weighted moving average (MEWMA) charts (Lowry et al. 1992), and the multivariate cumulative sum (MCUSUM) charts (Crosier 1998; Pignatiello and Runger 1990)) were constructed based on the assumption of normality. Even if observations do not comply with normality, with large samples, it seems to be not far-fetched to take an asymptotic normal approach to monitor the nonnormal data. However, as Stoumbos and Sullivan (2002) pointed out, when we have small samples from a nonnormal population, it is not reasonable to have the normality assumption by using the central limit theorem because the sample mean could be far from multivariate normal.

Recently, several techniques have been developed to construct multivariate control charts for nonnormal distributions. Especially, Hawkins and Maboudou Tchao (2008), Zou and Tsung (2008) and Reynolds and Stoumbos (2008) applied their methods to the MEWMA

control chart to monitor nonnormal observations.

To monitor exponential data from Gumbel bivariate exponential model (Gumbel 1960), Xie et al. (2011) proposed two bivariate exponentially weighted moving average (BEWMA) charts. With bivariate exponential data observed, Xie et al. (2011) considered three types of mean shifts:

(1) The D-D shift: the means of both random variables change downward;

(2) The U-U shift: the means of both random variables change upward;

(3) The U-D shift (D-U shift): the mean of one variable changes upward (downward) while the mean of the other changes downward (upward).

Among the three types of shifts, detecting the D-D shift is critical because we might experience the peculiar situation where the out-of-control ARL becomes larger than the in-control ARL. When detecting the D-D shift of exponential-type data, this phenomenon has been acknowledged by several researchers in the past. Stoumbos and Sullivan (2002) said "With a skewed distribution, such as the multivariate gamma, the shift direction affects the detection power. Furthermore, for some combinations of shift size and direction the OOC ARL exceeds the IC value. For a smaller set of shifts, the OOC ARL is even larger than the multi normal IC ARL, which is analogous to a biased test in hypothesis testing". Additionally, the recent work of Xie et al. (2011) clearly shows that their proposed control chart faces this odd situation.

In this chapter, we propose two monitoring methods that can be applied to univariate EWMA chart (Roberts 1959) in order to 1) prevent the abnormal results from happening in detecting the D-D shifts as well as 2) enhance the chart performance.

In Section 3.2, we introduce a copula function that is very useful to model bivariate joint distributions and propose two methods for their applications to the univariate EWMA chart. Next, extensive numerical comparisons to measure the effectiveness of our proposed methods are given in Section 3.3. The performance measure is the optimal ARL which is

the smallest ARL throughout the range of the smoothing parameter given a particular mean shift. In Section 3.4, with regard to the computation of the ARLs, we show an alternative way, which is the Markov chain method by using the empirical cumulative distribution function (CDF). The relationship between the chart performance and the correlation of the random variables is investigated in detail in Section 3.5. Finally, we discuss the advantages of the proposed methods and their potential extension to more than two random variables.

3.2 Construction of a Control Chart Using Bivariate Copulas

3.2.1 Gumbel Copulas

In this section, some basic concepts of copula functions are discussed and we demonstrate how multivariate joint distributions are constructed by using copula functions. A copula is a joint distribution generated from given marginal distributions. Unlike the multivariate normal distribution, it is very tricky to construct a multivariate joint distribution of skewed marginals when characteristics to be monitored are not independent. That is why the copula function becomes quite useful. The key benefit of copula functions is that they allow us to bind marginal distributions together and inject them varying degrees of correlation. Since we concentrate on only two quality variables, we look at the definition of a copula for the bivariate case. According to Sklar's theorem (Sklar 1959),

$C_\theta : [0,1]^2 \to [0,1]$ is a bivariate copula if the following conditions are satisfied:

(1) For every u, v in $[0,1], C_\theta(u,0) = 0 = C_\theta(0,v)$;

(2) For every u, v in $[0,1], C_\theta(u,1) = u$ and $C_\theta(1,v) = v$;

(3) For every u_1, u_2, v_1, v_2, in $[0,1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2, C_\theta(u_2,v_2) - C_\theta(u_2,v_1) - C_\theta(u_1,v_2) + C_\theta(u_1,v_1) \geq 0$. For every value of θ in some set Θ, θ is called the dependence parameter and its role is to produce varying degrees of the correlation between (X_1, X_2).

More importantly, the theorem states that with marginal CDFs $F_1(x)$ and $F_2(x)$ of a random vector (X_1, X_2), the function

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = C_\theta(F_1(x_1), F_2(x_2))$$

is a proper joint CDF and C_θ is a copula. Note that for the Gumbel copula, the dependence parameter is restricted on the region $[1, \infty)$. The Gumbel copula does not allow negative dependence. It also mentions that the copula is unique if each marginal function is continuous.

For the associated bivariate joint survival function $S(x_1, x_2)$, Nelsen (1999) showed that $S(x_1, x_2)$ can be expressed as follows:

$$S(x_1, x_2) = P(X_1 \geq x_1, X_2 \geq x_2) = \overline{C}_\theta(S_1(x_1), S_2(x_2)),$$

where $\overline{C}_\theta$ is given by

$$\overline{C}_\theta(u, v) = u + v - 1 + C_\theta(1 - u, 1 - v), \tag{3.1}$$

where $(u, v) \in [0, 1]^2$. For example, the Gumbel copula (Gumbel 1960) takes the form

$$C_\theta(u, v) = \exp\left\{-\left[(-\ln(u))^\theta + (-\ln(v))^\theta\right]^{1/\theta}\right\}, \tag{3.2}$$

where the dependence parameter θ lies in $[1, \infty)$. By using Equations (3.1) and (3.2), the survival copula associated with the Gumbel copula is then

$$\overline{C}_\theta(u, v) = u + v - 1 + \exp\left\{-\left[(-\ln(1 - u))^\theta + (-\ln(1 - v))^\theta\right]^{1/\theta}\right\}. \tag{3.3}$$

In this study, we use two bivariate joint distributions: 1) an Exponential-Exponential joint distribution and 2) a Gamma-Weibull joint distribution using the Gumbel copula.

As mentioned earlier, these joint distributions can be easily constructed by using copula functions. Throughout the chapter, among many copula functions, we choose to use the survival Gumbel copula (Equation (3.3)) to construct joint survival distributions of bivariate quality characteristics.

For example, consider $X_1 \sim \text{Exp}(\lambda_1)$ and $X_2 \sim \text{Exp}(\lambda_2)$. Then, by using Equation (3.3), the joint survival function of X_1 and X_2 can be constructed as follows:

$$\overline{F}_{X_1,X_2}(x_1,x_2) = \exp\left\{-\left[\left(\frac{x_1}{\lambda_1}\right)^{\theta} + \left(\frac{x_2}{\lambda_2}\right)^{\theta}\right]^{1/\theta}\right\}, \tag{3.4}$$

where $x_1, x_2 > 0,\ \lambda_1, \lambda_2 > 0,\ \theta \geq 1$.

If $X_1 \sim$ Gamma (α, β) and $X_2 \sim$ Weibull (k, λ), the joint survival distribution is

$$\overline{F}_{X_1,X_2}(x_1,x_2) = \exp\left\{-\left[\left(-\ln\left(1 - \frac{\gamma(\alpha, x_1/\beta)}{\Gamma(\alpha)}\right)\right)^{\theta} + \left(\frac{x_2}{\lambda}\right)^{k\theta}\right]^{1/\theta}\right\}, \tag{3.5}$$

where $x_1, x_2 > 0,\ \alpha, \beta, \lambda, k > 0,\ \theta \geq 1$ and $\gamma(\alpha, x_1/\beta) = \int_0^{x_1/\beta} y^{\alpha-1}\exp(-y)dy$ which is called the lower incomplete gamma function.

3.2.2 Bivariate EWMA Chart

Xie et al. (2011) proposed a new bivariate EWMA chart to monitor Gumbel's bivariate Exponential (Gumbel 1960) data. Following Lowry et al. (1992), the bivariate EWMA chart is given by

$$\mathbf{W}_t = r(\mathbf{X}_t - \boldsymbol{\mu}_0) + (1-r)\mathbf{W}_{t-1}, \quad t = 1, 2, \ldots,$$

where $\mathbf{X}_t = (X_{1t}, X_{2t})'$, r is the smoothing parameter $(0 < r \leq 1)$, $\mathbf{W}_0 = \mathbf{0}$ and $\boldsymbol{\mu}_0 = (\mu_{10}, \mu_{20})'$ is the on-target mean vector. The chart sends a signal when $[(2-r)/r]\,\mathbf{W}_t'\boldsymbol{\Sigma}_{\mathbf{X}_0}^{-1}\mathbf{W}_t > h$ by directly using the asymptotic in-control covariance matrix $\boldsymbol{\Sigma}_{\mathbf{W}} = [r/(2-r)]\boldsymbol{\Sigma}_{\mathbf{X}_0}$, where

$\mathbf{\Sigma}_{\mathbf{X}_0}$ is the in-control covariance matrix of $\mathbf{X}_t$ and h is the control limit. We refer to the procedure developed by Xie et al. (2011) as "Reference Method". We use the Reference Method for comparison purpose. In their study, they applied the Reference Method to the Gumbel bivariate Exponential (Gumbel 1960) distribution and relied on simulation to calculate the ARL values. In this chapter, by using the copula function (Equation (3.3)), we extend its application to the bivariate Gamma-Weibull distribution (Equation (3.5)) in order to handle more general distributions and make a more extensive comparison between our proposed methods and the Reference Method in the next section.

Suppose that we observe X_1 and X_2, the quality characteristics of interest to be monitored. We assume that the joint distribution of (X_1, X_2) follows either Equation (3.4) or (3.5). Keeping the main structure of the univariate EWMA chart (Roberts 1959), our proposed univariate EWMA statistic is defined as

$$Z_t = rP_t + (1-r)Z_{t-1}, \quad t = 1, 2, \ldots, \tag{3.6}$$

where r is the smoothing parameter $(0 < r \leq 1)$ and P_t is a function of two random variables X_1 and X_2. We take $Z_0 = E[P_t]$ calculated when the process is in-control. The chart issues a signal when either $Z_t > h$ or $Z_t < h$, depending on a function of P_t, where $h > 0$ is the control limit chosen to achieve a specified in-control ARL. We present two alternatives for P_t, each applied in conjunction with the univariate EWMA chart (Equation (3.6)) for monitoring the mean vector of (X_1, X_2).

Xie et al. (2011) considered three types of mean shifts to measure the chart performance: the D-D shift, U-U shift and U-D shift as described earlier. For a fair comparison, we use the three combinations of mean shifts mentioned before for the detection of the mean shifts.

3.2.3 Proposed Chart 1: The MAX-MIN Chart

With the three types of mean shifts considered, P_t is defined as one of the following schemes:

1. For the detection of the D-D shift, we consider $P_t = \min(X_{1t}, X_{2t})$ and the chart signals when $Z_t < h$, where $Z_0 = E\left[\min(X_{1t}, X_{2t})\right]$. Note that we use the lower control limit to detect the mean shift unlike the other cases;

2. For the detection of the U-U shift, we consider $P_t = \max(X_{1t}, X_{2t})$ and the chart signals when $Z_t > h$, where $Z_0 = E\left[\max(X_{1t}, X_{2t})\right]$;

3. For the detection of the U-D shift, we consider two statistics:

   (a) Let $P_t = \max(X_{1t}, X_{2t})/\min(X_{1t}, X_{2t})$. In this case, the chart signals when $Z_t > h$, where $Z_0 = E\left[\max(X_{1t}, X_{2t})/\min(X_{1t}, X_{2t})\right]$;

   (b) Let $P_t = \max(X_{1t}, X_{2t}) - \min(X_{1t}, X_{2t})$. In this case, the chart signals when $Z_t > h$, where $Z_0 = E\left[\max(X_{1t}, X_{2t}) - \min(X_{1t}, X_{2t})\right]$.

We call this approach as the "MAX-MIN Method".

3.2.4 Proposed Chart 2: The CDF Chart

This is an alternative approach to improve the chart performance. Here P_t is defined as one of the following schemes:

1. For the detection of the D-D shift, we consider $P_t = P(X_{1t} > x_{1t}, X_{2t} > x_{2t})$ and the chart signals when $Z_t > h$, where $Z_0 = E\left[P(X_{1t} > x_{1t}, X_{2t} > x_{2t})\right]$;

2. For the detection of the U-U shift, we consider $P_t = P(X_{1t} < x_{1t}, X_{2t} < x_{2t})$ and the chart signals when $Z_t > h$, where $Z_0 = E\left[P(X_{1t} < x_{1t}, X_{2t} < x_{2t})\right]$;

3. For the detection of the U-D shift, we consider $P_t = P(X_{1t} < x_{1t}, X_{2t} > x_{2t})$ and the chart signals when $Z_t > h$, where $Z_0 = E\left[P(X_t < x_{1t}, X_{2t} > x_{2t})\right]$.

Note that we use the upper control limit to detect the mean shift for all the cases, and call this approach as the "CDF Method". In either case, the expected values are calculated when the process is on-target.

3.3 Performance Comparions

To measure the performance of the proposed control charts, there are two areas that we need to investigate. Firstly, we need to check whether the design of our control charts prevents the out-of-control ARLs from becoming larger than the in-control ARLs for the detection of the D-D shifts. To show that our proposed methods are functioning properly, such an occurrence must not arise. Secondly, the ARL performance of the EWMA chart driven by either the MAX-MIN or the CDF Method has to be compared to that of the BEWMA control chart run by the Reference Method. Note that for performance comparisons, we use the optimal ARL, which is one of the most widely used performance measures for control charts.

In this section, for a bivariate joint distribution, we use the Exponential-Exponential and the Gamma-Weibull joint distributions, which correspond to the cases where 1) both random variables follow the same nonnormal distribution and 2) both variables follow completely different nonnormal distributions, respectively. Throughout this chapter, the numerical settings for both joint distributions are as follows.

For the Exponential-Exponential distribution, we will assume that when the process is on-target, $X_i \sim \exp(\lambda_i = 1), i = 1, 2$. By using Equation (3.4), the joint distribution of bivariate Exponential data (X_1, X_2) is given by

$$\overline{F}_{X_1,X_2}(x_1, x_2) = \exp\left\{-\left[(x_1)^\theta + (x_2)^\theta\right]^{1/\theta}\right\},$$

where $x_1, x_2 > 0,\ \theta \geq 1$. Thus, the in-control mean vector of the joint distribution is $\boldsymbol{\mu}_0 = (\mu_{10}, \mu_{20})' = (1, 1)'$ and the dependence parameter (θ) is 2. Note that the dependence

parameter ($\theta = 2$) corresponds to the correlation ($\rho = 0.57$) between X_1 and X_2. For the Exponential-Exponential joint distribution, given the dependence parameter, the correlation of X_1 and X_2 can be obtained by using the following equation (Lu and Bhattacharyya 1991a).

$$\rho = \frac{2\Gamma^2(1+1/\theta)}{\Gamma(2/\theta+1)} - 1.$$

For the Gamma-Weibull distribution, we will assume that $X_1 \sim$ Gamma ($\alpha = 3.660, \beta = 0.969$) and $X_2 \sim$ Weibull($k = 2, \lambda = 4$). The joint survival distribution is

$$\overline{F}_{X_1,X_2}(x_1,x_2) = \exp\left\{-\left[\left(-\ln\left(1-\frac{\gamma\left(3.660, x_1/0.969\right)}{\Gamma(3.660)}\right)\right)^{\theta} + \left(\frac{x_2}{4}\right)^{2\theta}\right]^{1/\theta}\right\},$$

where $x_1, x_2 > 0,\ \theta \geq 1$ and $\gamma(3.660, x/0.969) = \int_0^{x_1/0.969} y^{2.660}\exp(-y)dy$. The in-control mean of the joint distribution is $\boldsymbol{\mu}_0 = (\mu_{10}, \mu_{20})' = (3.55, 3.55)'$ and the dependence parameter θ is 2 as well. Especially, for the Gamma-Weibull joint distribution, since we are dealing with two-parameter distributions, we can assume that the variance of each random variable X_1 and X_2 remains constant after the process experiences a mean shift. Note that for this study, the in-control ARL (= ARL_0) is set at 200 and all the control limits are determined by one million simulations.

To measure the effectiveness of the control charts driven by our proposed methods, we use simulation of one million times to calculate the ARL values. In the next section we show that the ARL values can also be approximated by the Markov chain method by using the empirical CDF.

Regarding the simulation steps, here is the procedure how skewed data are generated for the simulation:

1. To generate (X_1, X_2), package 'VineCopula' in **R** (Schepsmeier et al. 2013) is used. For a fixed dependence parameter θ and $F_{X_i}(\cdot), i = 1, 2$, draw (u_1, u_2) from a copula function $C_\theta(u_1, u_2)$ using 'BiCopSim (p_1, p_2, p_3)', where the arguments p_1, p_2 and p_3

represent 1) number of observations, 2) copula family and 3) dependence parameter, respectively. For example, in order to generate 100 observations from the survival Gumbel copula (its family number is 14) with the dependence parameter ($\theta = 2$), then we write BiCopSim $(100, 14, 2)$;

2. Once (u_1, u_2) is generated from the copula function, X_1 and X_2 can be generated by using the quantile function of a probability function which is the inverse of its CDF. In other words, given (u_1, u_2), $x_1 = F_1^{-1}(u_1)$ and $x_2 = F_2^{-1}(u_2)$. For example, to generate X_1 and X_2 from the Exponential ($\lambda_1 = 1.2$)-Exponential ($\lambda_2 = 0.7$) joint distribution, we generate X_1 and X_2 by writing $x_1 = \text{qexp}\ (u_1, 1.2)$ and $x_2 = \text{qexp}\ (u_2, 0.7)$;

3. Calculate either the MAX-MIN or the CDF statistics for X_1 and X_2;

4. Finally, obtain the univariate EWMA statistic (Equation (3.6)).

Tables 3.1 and 3.2 present the ARLs for various smoothing parameters in order to detect D-D shifts of the Exponential-Exponetial and the Gamma-Weibull joint distributions. The tables clearly show that the EWMA charts driven by our proposed methods do not show any sign of the "undesirable phenomenon" unlike the chart of the Reference Method. For example, from Table 3.1, suppose that we choose 0.1 as the smoothing parameter and $\boldsymbol{\mu}_0 = (\mu_{10}, \mu_{20})' = (1.0, 1.0)'$ has shifted to $\boldsymbol{\mu}_1 = (\mu_{11}, \mu_{21})' = (0.8, 0.8)'$. In this setting, the control chart of the Reference Method provides 1583.81 as the out-of-control ARL which is a completely illogical number. On the other hand, the EWMA chart driven by either the MAX-MIN or the CDF Method comes up with valid off-target ARLs. They are all smaller than the in-control ARL which is 200.

Moreover, from other numerical evidence, it gets worse as the departure from the in-control means gets smaller. In others words, upon reviewing the example of Table 3.1, the chart of the Reference Method will not function properly unless we use a smaller smoothing parameter than 0.05 to detect $\boldsymbol{\mu}_1 = (\mu_{11}, \mu_{21})' = (0.9, 0.9)'$ properly. Therefore, the chart

of the Reference Method provides only a fraction of smoothing parameters to be chosen for the meaningful out-of-control ARL values. Conventionally, the use of the smaller smoothing parameter gives an advantage to the EWMA-type control charts over other types of charts in monitoring small mean shifts. However, from another point of view, being forced to use a smaller smoothing parameter makes the chart more exposed to the adverse effect of so-called "chart inertia" (Woodall and Mahmoud 2005). On the other hand, the EWMA chart driven by our proposed methods works in a stable way, providing meaningful out-of-control ARLs.

To examine whether our new approach enhances the chart performance, the optimal ARL is measured for each possible mean shift: the D-D, U-U and U-D shifts. Tables 3.4 - 3.9 present the comparisons of the optimal ARLs between methods and the efficiencies of our new approach with respect to the Reference Method for given mean shifts. Note that numbers in parentheses are the optimal smoothing parameters. To measure the efficiency of our new approach, we use the following quantity:

$$\text{Efficiency} = 100\frac{\text{ARL}^{(N)} - \text{ARL}^{(O)}}{\text{ARL}^{(O)}}\%,$$

where $\text{ARL}^{(N)}$ and $\text{ARL}^{(O)}$ are the optimal average run lengths from the new and old methods, respectively. Thus, this is really a measure of relative efficiency.

3.3.1 Detection of D-D Shifts

Examination of Tables 3.4 and 3.5 reveals that the EWMA chart run by the MAX-MIN Method outperforms the bivariate EWMA chart run by the Reference Method by a large margin in every direction used. Whether comparisons are conducted with the Exponential-Exponential joint distribution or the Gamma-Weibull joint distribution, it is clear that the MAX-MIN Method presents a substantial reduction in the ARL_{opt} for the detection of D-D shifts. Based on the directions used in Tables 3.4 and 3.5, with the MAX-MIN Method applied to the EWMA chart, 44.6% and 42.9% of improvement relative to the Reference

Method were recorded, respectively. On the other hand, the chart run by the CDF Method also performs better than the chart of the Reference Method in many directions but there are some directions where the CDF Method does not perform well. The results show that if the method is used in monitoring a mean shift where only one variable changes, the chart run by the CDF Method performs poorly. To make the comparison of the chart performance clearer, optimal ARLs of each method are plotted in Figure 3.1. The x-axis represents the directions described in Tables 3.4 and 3.6. Note that the spiky points observed in Figure 3.1 indicate the area where the CDF Method performs poorly. To address this issue, we suggest using just one variable with the same methods since only one variable is a matter of our concern. Table 3.3 clearly shows that the CDF Method with one variable improves the chart performance in every direction where the CDF Method with two variables did not perform well.

3.3.2 Detection of U-U Shifts

As can be seen from Tables 3.6 and 3.7, both the MAX-MIN and the CDF methods outperform the Reference Method in many directions. However, both methods struggle to show competitive performance to the Reference Method in the areas where only one variable changes. The same issue that the CDF Method has for the detection of D-D shifts appears in both methods. More distinctive spiky points indicating poor chart performance of our methods are observed in Figure 3.1(b). Notice that the CDF Method is more vulnerable to those directions than the MAX-MIN Method. To tackle this problem, just as for the detection of D-D shifts, either the MAX-MIN or the CDF Method with one variable is recommended for use.

Table 3.1: Comparisons of ARLs for D-D shifts of Exponential-Exponential Joint Distribution When $\delta = 2$ and $\text{ARL}_0 = 200$

The Reference Method							
(μ_{11}, μ_{21})	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	7.4185	10.3501	14.7517	18.2900	21.3000	23.3980
(0.9, 0.9)		$*$	$*$	$*$	$*$	$*$	$*$
(0.8, 0.8)		179.09	$*$	$*$	$*$	$*$	$*$
(0.5, 0.5)		27.76	110.48	$*$	$*$	$*$	$*$
(0.9, 0.6)		40.08	144.92	$*$	$*$	$*$	$*$
(0.5, 0.2)		15.65	23.01	$*$	$*$	$*$	$*$
(1.0, 0.8)		94.63	186.21	$*$	$*$	$*$	$*$
The MAX-MIN Method							
(μ_{11}, μ_{21})	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	0.5264	04268	0.3065	0.2305	0.1756	0.1331
(0.9, 0.9)		85.19	95.58	109.99	118.83	125.78	133.26
(0.8, 0.8)		44.68	50.59	61.58	70.59	78.72	87.27
(0.5, 0.5)		14.38	13.92	14.71	16.51	19.04	22.36
(0.9, 0.6)		28.33	30.85	37.34	44.10	50.80	58.28
(0.5, 0.2)		8.89	7.86	7.01	6.79	6.97	7.46
(1.0, 0.8)		77.26	87.33	101.81	111.12	118.65	126.70
The CDF Method							
(μ_{11}, μ_{21})	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	0.4602	0.5196	0.6084	0.6783	0.7370	0.7877
(0.9, 0.9)		92.34	103.52	118.84	129.60	135.73	140.82
(0.8, 0.8)		48.95	56.28	70.07	80.96	89.80	98.43
(0.5, 0.5)		12.94	13.32	15.69	18.87	22.66	27.17
(0.9, 0.6)		37.76	43.38	55.34	65.10	74.58	82.78
(0.5, 0.2)		8.51	8.25	8.65	9.62	11.06	13.09
(1.0, 0.8)		93.51	103.58	120.40	130.11	134.70	141.69

$*$ indicates that the off-target ARL is bigger than the on-target ARL which is 200.

3.3.3 Detection of U-D Shifts

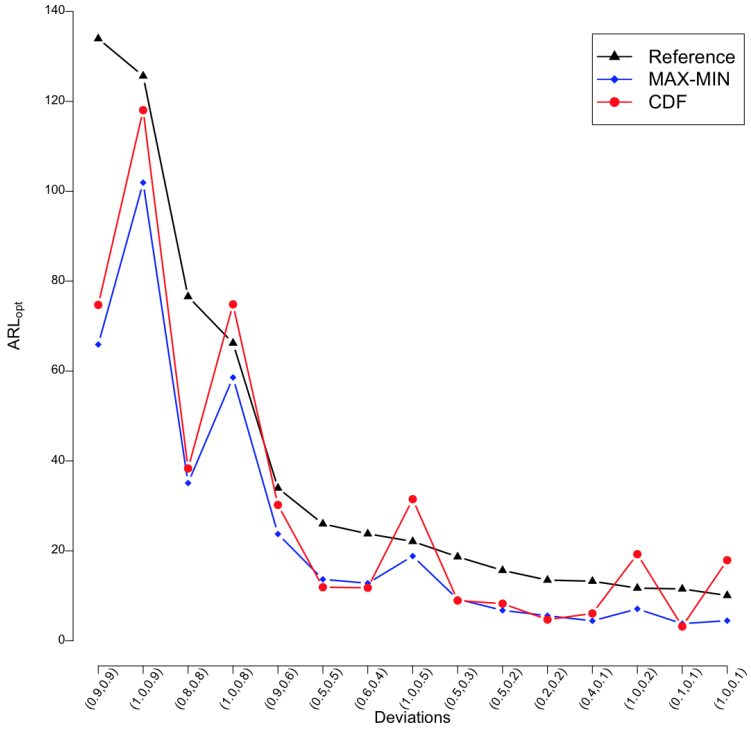
Tables 3.8 and 3.9 display that the CDF Method outperforms other methods in all directions used for both the Exponential-Exponential joint distribution and the Gamma-Weibull joint distributions. Based on all the directions used, the CDF Method improves the chart performance by 38.8% and 35.3%, respectively, for the two distributions. For the detection of U-D shifts of the Exponential-Exponential joint distribution, clearly the CDF Method is the best choice but the MAX-MIN Method does not provide good performance consistently. However,

Table 3.2: Comparisons of ARLs for D-D Shifts of Gamma-Weibull Joint Distribution When $\delta = 2$ and $\text{ARL}_0 = 200$

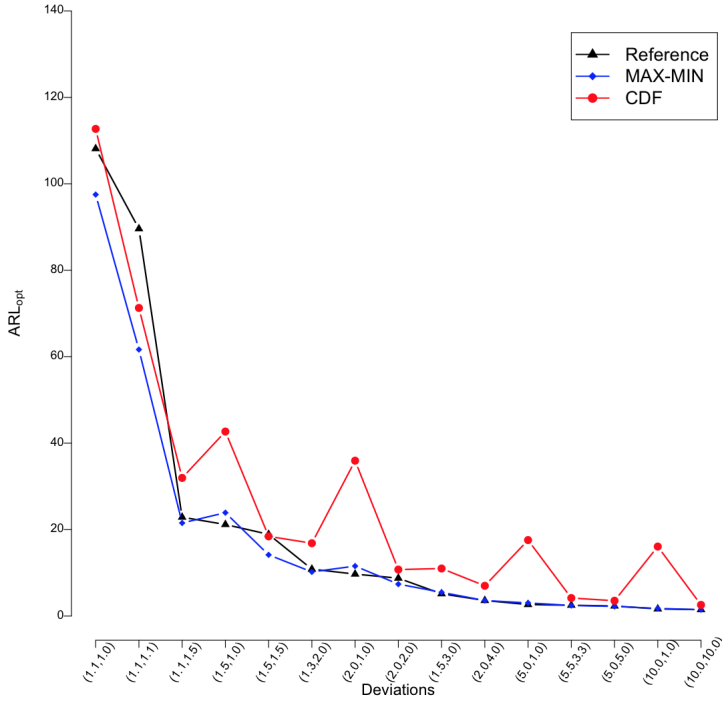
The Reference Method							
(μ_{11}, μ_{21})	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	7.3860	8.9775	11.0731	12.6650	13.9281	14.9423
(3.5, 3.5)		$*$	$*$	$*$	$*$	$*$	$*$
(3.0, 3.0)		48.27	66.03	199.26	$*$	$*$	$*$
(2.5, 2.5)		19.53	20.60	40.28	149.0	179.73	$*$
(3.5, 3.55)		197.59	$*$	$*$	$*$	$*$	$*$
(3.2, 3.55)		69.99	97.71	169.70	$*$	$*$	$*$
The MAX-MIN Method							
(μ_{11}, μ_{21})	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	2.5559	2.2931	1.9281	1.6539	1.4275	1.2283
(3.5, 3.5)		153.12	155.82	162.45	164.61	168.61	170.56
(3.0, 3.0)		22.89	23.37	24.89	26.93	29.04	31.48
(2.5, 2.5)		10.64	9.94	9.41	9.39	9.57	9.93
(3.5, 3.55)		171.83	175.56	178.59	182.37	184.02	185.79
(3.2, 3.55)		67.34	72.97	81.16	89.04	96.10	103.54
The CDF Method							
(μ_{11}, μ_{21})	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	0.4603	0.5198	0.6086	0.6785	0.7371	0.7875
(3.5, 3.5)		158.33	163.10	165.33	168.62	174.13	172.91
(3.0, 3.0)		26.21	28.12	31.46	34.47	37.04	39.09
(2.5, 2.5)		11.92	11.80	12.06	12.61	13.10	13.63
(3.5, 3.55)		173.47	177.55	180.39	182.36	183.77	184.05
(3.2, 3.55)		85.00	92.69	120.40	107.34	111.27	114.34

$*$ indicates that the off-target ARL is bigger than the on-target ARL which is 200.

when the Gamma-Weibull joint distribution is used, the MAX-MIN Method also shows competitiveness in all directions investigated, which would be an alternative choice in this case. Note that, as discussed in Section 3.2.3, we used two statistics for the MAX-MIN Method for the detection of U-D shifts. Regarding the chart performance, $\max(X_1, X_2)/\min(X_1, X_2)$ seems to provide better performance than $\max(X_1, X_2) - \min(X_1, X_2)$.



(a)



(b)

Figure 3.1: Chart Performances of Exponential-Exponential Distribution (a) D-D Shifts; (b) U-U Shifts.

Table 3.3: Performance of CDF Method with One Variable for D-D Shifts

Exponential-Exponential Distribution	CDF Method	Reference Method
Deviation	ARL_{opt}	ARL_{opt}
0.9	**73.52**(0.01)	125.65 (0.01)
0.8	**38.42**(0.01)	66.22 (0.01)
0.5	**12.61**(0.01)	22.09 (0.05)
0.2	**5.26**(0.31)	11.71 (0.05)
0.1	**3.59**(0.55)	10.07 (0.07)
Gamma-Weibull Distribution	CDF Method	Reference Method
Deviation	ARL_{opt}	ARL_{opt}
3.5	**137.15**(0.01)	187.73 (0.01)
3.0	**21.97**(0.01)	37.73 (0.01)
2.5	**11.44** (0.01)	15.84 (0.05)
2.0	**5.77** (0.01)	9.07 (0.05)
1.5	**3.66** (0.31)	6.23 (0.07)
1.0	**2.73** (0.65)	4.53 (0.07)

- The values in bold indicate the best performance among methods.

Table 3.4: The Optimal Out-of-Control ARLs for D-D Shifts of Exponential-Exponential Joint Distribution

D-D Shift	MAX-MIN Method		CDF Method		Reference Method
Deviation	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}
(0.9, 0.9)	**65.90** (0.01)	-50.8%	74.72 (0.01)	-44.2%	133.95 (0.001)
(0.8, 0.8)	**35.08** (0.01)	-54.2%	33.33 (0.01)	-49.9%	76.57 (0.01)
(0.5, 0.5)	13.66 (0.01)	-47.5%	**11.90** (0.01)	-54.2%	26.00 (0.02)
(0.2, 0.2)	5.57 (0.40)	-58.7%	**4.70** (0.22)	-65.2%	13.49 (0.05)
(0.1, 0.1)	3.80 (0.55)	-67.0%	**3.17** (0.49)	-72.5%	11.52 (0.07)
(0.9, 0.6)	**23.74** (0.01)	-30.1%	30.22 (0.01)	-11.0%	33.97 (0.02)
(0.5 0.2)	**6.76** (0.31)	-56.8%	8.25 (0.10)	-47.3%	15.65 (0.05)
(0.4, 0.1)	**4.42** (0.51)	-66.6%	6.07 (0.20)	-54.2%	13.24 (0.05)
(0.6, 0.4)	12.79 (0.10)	-46.2%	**11.78** (0.01)	-50.5%	23.79 (0.02)
(0.5, 0.3)	9.25 (0.20)	-50.5%	**8.94** (0.01)	-52.1%	18.67 (0.05)
(1.0, 0.9)	**101.92** (0.01)	-18.9%	118.05 (0.01)	-6.0%	125.65 (0.01)
(1.0, 0.8)	**58.57** (0.01)	-11.6%	74.84 (0.01)	13.0%	66.22 (0.01)
(1.0, 0.5)	**18.83** (0.01)	-14.8%	31.50 (0.01)	42.6%	22.09 (0.05)
(1.0, 0.2)	**7.08** (0.30)	-39.5%	19.25 (0.01)	64.4%	11.71 (0.05)
(1.0, 0.1)	**4.46** (0.50)	-55.7%	17.91 (0.01)	77.9%	10.07 (0.07)

- The values in bold indicate the best performance among methods.

3.4 Verification of ARLs

In the previous sections, ARLs were calculated by using a large number of simulations and the accuracy of the numbers has to be checked. To verify the accuracy of the ARLs obtained

Table 3.5: The Optimal Out-of-Control ARLs for D-D Shifts of Gamma-Weibull Joint Distribution

D-D Shift	MAX-MIN Method		CDF Method		Reference Method
Deviation	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}
(3.5, 3.5)	**138.32** (0.01)	-27.9%	142.56 (0.01)	-25.7%	191.77 (0.01)
(3.0, 3.0)	**20.14** (0.01)	-56.3%	21.88 (0.01)	-52.5%	46.11 (0.02)
(2.5, 2.5)	**9.39** (0.30)	-51.9%	10.95 (0.01)	-43.9%	19.53 (0.05)
(2.0, 2.0)	**4.87** (0.50)	-56.9%	6.76 (0.40)	-40.1%	11.29 (0.10)
(1.5, 1.5)	**3.08** (0.60)	-59.3%	4.21 (0.50)	-44.4%	7.57 (0.10)
(1.0, 1.0)	**2.36** (0.65)	-56.6%	3.01 (0.65)	-44.7%	5.44 (0.20)
(3.0, 2.5)	**12.72** (0.20)	-41.5%	15.12 (0.01)	-30.5%	21.76 (0.05)
(2.0, 1.0)	**2.70** (0.67)	-57.5%	3.01 (0.65)	-52.7%	6.36 (0.20)
(1.5, 2.0)	**3.55** (0.61)	-58.4%	5.55 (0.53)	-35.1%	8.55 (0.10)
(1.0, 1.5)	**2.54** (0.72)	-58.6%	3.66 (0.55)	-40.4%	6.14 (0.21)
(2.0, 1.5)	**3.60** (0.60)	-57.7%	5.32 (0.49)	-37.5%	8.51 (0.11)
(2.5, 2.0)	**6.16** (0.52)	-52.4%	9.08 (0.21)	-29.8%	12.93 (0.11)
(3.5, 3.54)	**160.69** (0.01)	-14.4%	163.77 (0.01)	-12.8%	187.73 (0.01)
(3.0, 3.54)	**33.67** (0.01)	-10.7%	47.95 (0.01)	27.1%	37.73 (0.05)
(2.5, 3.54)	**15.45** (0.01)	-2.5%	29.43 (0.01)	85.7%	15.84 (0.05)
(2.0, 3.54)	**7.37** (0.50)	-18.7%	23.01 (0.01)	153.8%	9.07 (0.10)
(1.5, 3.54)	**4.08** (0.64)	-34.5%	20.17 (0.01)	223.7%	6.23 (0.20)
(1.0, 3.54)	**2.78** (0.71)	-38.6%	18.64 (0.01)	311.1%	4.53 (0.20)

- The values in bold indicate the best performance among methods.

Table 3.6: The Optimal Out-of-Control ARLs for U-U Shifts of Exponential-Exponential Joint Distribution

U-U Shift	MAX-MIN Method		CDF Method		Reference Method
Deviation	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}
(1.1, 1.1)	**61.66** (0.01)	-31.2%	71.29 (0.01)	-20.4%	89.61 (0.07)
(1.5, 1.5)	**14.16** (0.01)	-25.1%	18.44 (0.01)	-2.4%	18.90 (0.07)
(2.0, 2.0)	**7.39** (0.01)	-15.6%	10.75 (0.01)	-22.7%	8.76 (0.20)
(5.0, 5.0)	**2.23** (0.40)	-3.9%	3.54 (0.65)	-52.6%	2.32 (0.40)
(10.0, 10.0)	**1.47** (0.50)	-0.7%	2.58 (0.64)	74.3%	1.48 (0.60)
(1.3, 2.0)	**10.21** (0.01)	-5.9%	16.85 (0.01)	55.3%	10.85 (0.10)
(5.5, 3.3)	**2.44** (0.40)	-2.3%	4.17 (0.51)	66.8%	2.50 (0.52)
(1.5, 3.0)	5.22 (0.20)	-6.6%	11.00 (0.01)	113.6%	**5.21** (0.20)
(1.1, 1.5)	**21.53** (0.01)	-5.9%	31.96 (0.01)	39.7%	22.87 (0.07)
(2.0, 4.0)	**3.59** (0.20)	-0.2%	7.00 (0.40)	94.7%	3.60 (0.30)
(1.1, 1.0)	**97.52** (0.01)	-9.8%	112.72 (0.01)	4.2%	108.14 (0.07)
(1.5, 1.0)	23.92 (0.01)	12.9%	42.68 (0.01)	101.5%	**21.18** (0.05)
(2.0, 1.0)	11.56 (0.01)	19.1%	35.95 (0.01)	270.4%	**9.70** (0.10)
(5.0, 1.0)	3.03 (0.40)	13.9%	17.57 (0.01)	560.6%	**2.66**(0.30)
(10.0, 1.0)	1.76 (0.70)	6.9%	16.08 (0.01)	876.6%	**1.65** (0.30)

- The values in bold indicate the best performance among methods.

by simulation, we use the Markov chain method. To apply the Markov chain method to the control chart (Equation (3.6)), the distribution of P_t has to be known, i.e., it has to be

Table 3.7: The Optimal Out-of-Control ARLs for U-U Shifts of Gamma-Weibull Joint Distribution

U-U Shift	MAX-MIN Method		CDF Method		Reference Method
Deviation	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}
(3.6, 3.6)	136.37 (0.01)	-24.6%	**122.55** (0.01)	-32.2%	180.83 (0.10)
(4.0, 4.0)	32.16 (0.01)	-43.5%	**24.80**(0.01)	-56.5%	56.97 (0.02)
(4.5, 4.5)	15.26 (0.01)	-32.7%	**11.87** (0.01)	-47.6%	22.67 (0.05)
(5.0, 5.0)	9.96 (0.01)	-22.8%	**7.96** (0.01)	-38.3%	12.90 (0.10)
(6.0, 6.0)	5.85 (0.11)	-8.6%	**4.15** (0.40)	-35.2%	6.40 (0.20)
(8.0, 8.0)	2.69 (0.30)	-6.6%	**2.25** (0.50)	-21.9%	2.88 (0.30)
(4.5, 5.0)	11.64 (0.01)	-20.1%	**9.69** (0.01)	-33.5%	14.57 (0.10)
(5.0, 6.0)	6.99 (0.10)	-2.9%	**5.84** (0.20)	-18.9%	7.20 (0.10)
(3.7, 4.0)	44.93 (0.01)	-22.0%	**38.30** (0.01)	-33.5%	57.58 (0.02)
(4.0, 4.5)	19.58 (0.01)	-20.9%	**16.78** (0.01)	-32.2%	24.73 (0.05)
(6.5, 8.0)	3.04 (0.30)	-6.5%	**2.73** (0.50)	-16.0%	3.25 (0.30)
(6.0, 8.0)	**3.10** (0.31)	-3.2%	3.13 (0.52)	-2.5%	3.21 (0.30)
(3.6, 3.54)	165.04 (0.01)	-10.3%	**155.09** (0.01)	-15.7%	184.00 (0.10)
(4.0, 3.54)	54.43 (0.01)	13.8%	53.20 (0.01)	11.3%	**47.82** (0.02)
(4.5, 3.54)	23.90 (0.01)	29.4%	30.39 (0.01)	64.6%	**18.47** (0.05)
(5.0, 3.54)	14.26 (0.01)	38.2%	23.14 (0.01)	124.2%	**10.32** (0.10)
(6.0, 3.54)	7.53 (0.01)	48.4%	18.12 (0.01)	257.3%	**5.07** (0.20)
(8.0, 3.54)	3.24 (0.30)	39.4%	15.95 (0.01)	586.3%	**2.32** (0.40)

- The values in bold indicate the best performance among methods.

Table 3.8: The Optimal Out-of-Control ARLs for U-D Shifts of Exponential-Exponential Joint Distribution

U-D Shift	MAX-MIN Method				CDF Method		Reference Method
	$\max(X_1, X_2)/\min(X_1, X_2)$		$\max(X_1, X_2) - \min(X_1, X_2)$				
Deviation	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}
(1.5, 0.8)	30.96 (0.01)	89.4%	19.43 (0.01)	18.9%	**10.67** (0.01)	-34.7%	16.34 (0.05)
(2.0, 0.5)	5.75 (0.01)	-17.3 %	6.74(0.01)	-3.2%	**4.40** (0.01)	-36.8%	6.96 (0.07)
(1.5, 0.5)	9.33 (0.01)	-14.8 %	12.84 (0.01)	17.3%	**5.70** (0.01)	-47.9%	10.95 (0.05)
(5.0, 0.2)	1.29 (0.01)	-46.5%	2.13 (0.40)	-11.6%	**1.52** (0.50)	-36.9%	2.41 (0.40)
(10 0.1)	1.02 (0.01)	-35.4%	1.46 (0.50)	-7.0%	**1.14** (0.60))	-27.2%	1.58 (0.50)
(1.1, 0.9)	148.76 (0.01)	101.0 %	134.40 (0.01)	81.6%	**36.95** (0.01)	-50.1%	74.02 (0.01)
(1.3, 0.6)	19.85 (0.01)	19.7 %	25.57 (0.01)	54.3%	**8.76** (0.01)	-47.1%	16.58 (0.05)
(3.5, 0.5)	2.96 (0.01)	-13.6 %	3.18 (0.19)	-7.3%	**2.94** (0.30)	-14.3%	3.43 (0.21)
(1.4, 0.3)	4.67 (0.01)	-50.5 %	11.16 (0.01)	41.4%	**4.14** (0.01)	-56.1%	9.42 (0.07)
(2.5, 0.4)	3.31 (0.01)	-32.2 %	4.49 (0.07)	-8.2%	**3.08** (0.28)	-37.0%	4.89 (0.10)

- The values in bold indicate the best performance among methods.

possible to compute it numerically. Regarding the correlation between two random variables X_1 and X_2, there are two possible cases: 1) independent case and 2) dependent case.

Table 3.9: The Optimal Out-of-Control ARLs for U-D Shifts of Gamma-Weibull Joint Distribution

U-D Shift	MAX-MIN Method				CDF Method		Reference Method
	$\max(X_1, X_2)/\min(X_1, X_2)$		$\max(X_1, X_2) - \min(X_1, X_2)$				
Deviation	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}	Efficiency	ARL_{opt}
(4.0, 3.0)	12.39 (0.01)	-37.0%	26.02 (0.01)	32.2%	**10.05** (0.01)	-48.9%	19.68 (0.05)
(6.0, 2.0)	1.91 (0.31)	-35.8%	2.69 (0.20)	-9.4%	**1.83**(0.50)	-38.4%	2.97 (0.30)
(9.0, 1.0)	1.14 (0.55)	-0.9 %	1.06 (0.70)	-7.8%	**1.13** (0.40)	-1.7%	1.15 (0.70)
(5.0, 3.0)	5.89 (0.01)	-20.7%	7.82 (0.01)	5.2%	**4.79** (0.20)	-35.5%	7.43 (0.10)
(5.5, 2.5)	2.93 (0.01)	-31.5%	4.28 (0.09)	0.0%	**2.74** (0.30)	-36.0%	4.28 (0.20)
(6.0, 3.0)	3.92 (0.01)	-5.8 %	4.33 (0.11)	4.1%	**3.09** (0.29)	-25.7%	4.16 (0.21)
(4.5, 3.2)	11.43 (0.01)	-14.8 %	16.29 (0.01)	21.5%	**7.54** (0.01)	-43.8%	13.41 (0.10)
(3.9, 3.1)	17.71 (0.01)	-32.8 %	38.93 (0.10)	47.8%	**13.01** (0.01)	-50.6%	26.34 (0.05)
(5.5, 1.0)	1.53 (0.44)	-40.2%	2.23 (0.30)	-12.9%	**1.48** (0.61)	-42.2%	2.56 (0.31)
(5.5, 3.2)	5.82 (0.01)	-3.2 %	6.31 (0.01)	5.0%	**4.22** (0.21)	-29.8%	6.01 (0.20)

- The values in bold indicate the best performance among methods.

3.4.1 Exact Distribution of Chart Statistic - CDF Method

When the random variables X_1 and X_2 are independent, it is possible to derive the distribution of P_t analytically. For example, the distribution of $P_t = P(X_{1t} < x_{1t}, X_{2t} < x_{2t}, \ldots, X_{pt} < x_{2t})$, which is used for the detection of U-U shifts can be derived in the following way.

Consider p continuous and independent random variables $X_1, X_2, \ldots, X_p$. Assume that X_i has CDF $F_i(x_i), i = 1, 2, \ldots, p$. Then the joint CDF of $X_1, X_2, \ldots X_p$ is given by

$$\begin{aligned} F(x_1, x_2, \ldots, x_p) &= P\{X_1 \le x_1, X_2 \le x_2, \ldots, X_p \le x_p\} \\ &= F_1(x_1)F_2(x_2)\cdots F_p(x_p) \\ &= \prod_{i=1}^{p} F_i(x_i)\ . \end{aligned}$$

In order to obtain the distribution of $F(\mathbf{X})$ where $\mathbf{X} = (X_1, X_2, \ldots, X_p)'$, let us define $Z = F(\mathbf{X})$ and $Z_i = F_i(\mathbf{X}_i) \sim U(0, 1)$ for $i = 1, 2, \ldots, p$. It is clear that $Z = \prod_{i=1}^{p} Z_i$ and that $Z_1, Z_2, \ldots, Z_p$ are independent and identically distributed $U(0, 1)$ random variables.

Thus, the probability density function of each Z_i is

$$f_i(z_i) = \begin{cases} 1 & 0 < z_i < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

Now let us consider the transformation $Y = -\ln(Z)$. Clearly, $Y = \sum_{i=1}^{p} Y_i$ and $Y_i = -\ln(Z_i), i = 1, 2, \ldots, p$.

In order to find the distribution of Y, consider the transformation $y_i = -\ln(z_i), z_i = e^{-y_i}$, and $\dfrac{dz_i}{dy_i} = -e^{-y_i}$. By the univariate transformation formula in conjunction with Equation (3.7), we get

$$f_{Y_i}(y_i) = f_{Z_i}(z_i = e^{-y_i}) \left| \frac{dz_i}{dy_i} \right| = e^{-y_i},\ 0 < y_i < \infty.$$

Thus, $Y_i \sim$ Exp $(\lambda = 1)$. In other words, $Y_i \sim$ Gamma $(\alpha = 1, \beta = 1)$. By the additivity property of independent Gamma distributed random variables, we get

$$Y \sim \text{Gamma } (\alpha = p, \beta = 1).$$

Therefore, the probability density function of Y is

$$f_Y(y) = \begin{cases} \dfrac{y^{p-1}e^{-y}}{\Gamma(p)} & 0 < y < \infty \\ 0 & \text{otherwise.} \end{cases} \tag{3.8}$$

Let us transform back from Y to Z to obtain the distribution of Z:

$$y = -\ln(z),\ \frac{dy}{dz} = -\frac{1}{z}.$$

Thus, from Equation (3.8) we get the probability density function of Z as

$$f_Z(z) = \begin{cases} \dfrac{[-\ln(z)]^{p-1}}{(p-1)!} & 0 < z < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Note that this is a monotonic decreasing function of z for every $p \geq 2$. As for $p = 2$, if X_1 and X_2 are independent continuous random variables, then their joint distribution of Z has a distribution with probability density function given by

$$f_Z(z) = \begin{cases} -\ln(z) & 0 < z < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Accordingly, the cumulative distribution function of $Z = P(X_1 < x_1, X_2 < x_2)$ is

$$F_Z(z) = z - z\ln(z), \quad 0 < z < 1. \tag{3.9}$$

Note that the distributions of $P(X_{1t} > x_{1t}, X_{2t} > x_{2t})$, $P(X_{1t} > x_{1t}, X_{2t} < x_{2t})$ and $P(X_{1t} < x_{1t}, X_{2t} > x_{2t})$ all have the same distribution as $P(X_{1t} < x_{1t}, X_{2t} < x_{2t})$ and $X_{it}, i = 1, 2$ are the quality variables at time t.

3.4.2 Exact Distribution of Chart Statistic - MAX-MIN Method

Furthermore, the distributions of $P_t = \max(X_{1t}, X_{2t})/\min(X_{1t}, X_{2t})$ and $P_t = \max(X_{1t}, X_{2t}) - \min(X_{1t}, X_{2t})$ which are used for the detection of U-D shifts can be derived as well. Consider two continuous and independent random variables and $X_i \sim \text{Exp}(\theta_i)$ where $i = 1, 2$. Let $X_{(1)}$ and $X_{(2)}$ denote the order statistics. Clearly, $\max(X_1, X_2) = X_{(2)}$ and $\min(X_1, X_2) = X_{(1)}$. Let us define that $V = X_{(2)}/X_{(1)}$ and $R = X_{(2)} - X_{(1)}$. We will derive the joint probability density function of R and V from the joint probability density function of $X_{(1)}$ and $X_{(2)}$.

The joint probability density function of $X_{(1)}$ and $X_{(2)}$ is

$$f_{X_{(1)},X_{(2)}}(x_1,x_2) = \frac{1}{\theta_1\theta_2}\left(e^{(-x_1/\theta_1)}e^{(-x_2/\theta_2)} + e^{(-x_1/\theta_2)}e^{(-x_2/\theta_1)}\right),\ 0 < x_1 < x_2 < \infty. \quad (3.10)$$

Solving for $X_{(1)}$ and $X_{(2)}$, we obtain $X_{(1)} = \dfrac{R}{V-1}$ and $X_{(2)} = \dfrac{VR}{V-1}$. The Jacobian of this transformation is given by

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial r} & \dfrac{\partial x_1}{\partial v} \\ \dfrac{\partial x_2}{\partial r} & \dfrac{\partial x_2}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{v-1} & \dfrac{-r}{(v-1)^2} \\ \dfrac{v}{(v-1)} & \dfrac{-r}{(v-1)^2} \end{vmatrix} = \frac{r}{(v-1)^2}.$$

Thus, from Equation (3.10) we obtain the joint probability function as

$$\begin{aligned} f_{R,V}(r,v) &= \frac{1}{\theta_1\theta_2}\left\{\exp\left(\frac{-r}{\theta_1(v-1)}\right)\exp\left(\frac{-vr}{\theta_2(v-1)}\right) + \exp\left(\frac{-r}{\theta_2(v-1)}\right)\exp\left(\frac{-vr}{\theta_1(v-1)}\right)\right\} \\ &\times \frac{r}{(v-1)^2}, \quad 0 < r < \infty, \quad 1 < v < \infty. \end{aligned} \quad (3.11)$$

By integrating out R in Equation (3.11), the marginal probability density function of V is obtained,

$$\begin{aligned} f_V(v) &= \frac{1}{\theta_1\theta_2}\int_0^\infty \exp\left(\frac{-r}{\theta_1(v-1)}\right)\exp\left(\frac{-vr}{\theta_2(v-1)}\right)\frac{r}{(v-1)^2}dr \\ &+ \frac{1}{\theta_1\theta_2}\int_0^\infty \exp\left(\frac{-r}{\theta_2(v-1)}\right)\exp\left(\frac{-vr}{\theta_1(v-1)}\right)\frac{r}{(v-1)^2}dr, \\ &= \frac{\theta_1\theta_2}{(\theta_1 v+\theta_2)^2} + \frac{\theta_1\theta_2}{(\theta_2 v+\theta_1)^2}, \quad 1 < v < \infty. \end{aligned}$$

The cumulative distribution function of $V = \max(X_1,X_2)/\min(X_1,X_2)$ is

$$F_V(v) = 1 - \frac{\theta_1}{(\theta_2 v+\theta_1)} - \frac{\theta_2}{(\theta_1 v+\theta_2)}, \quad 1 < v < \infty. \quad (3.12)$$

Additionally, the marginal probability density function of R is given by

$$\begin{aligned} f_R(r) &= \frac{1}{\theta_1\theta_2}\int_1^\infty \exp\left(\frac{-r}{\theta_1(v-1)}\right)\exp\left(\frac{-vr}{\theta_2(v-1)}\right)\frac{r}{(v-1)^2}dv \\ &+ \frac{1}{\theta_1\theta_2}\int_1^\infty \exp\left(\frac{-r}{\theta_2(v-1)}\right)\exp\left(\frac{-vr}{\theta_1(v-1)}\right)\frac{r}{(v-1)^2}dv, \\ &= \frac{\exp\left(\frac{-r}{\theta_2}\right)}{(\theta_1+\theta_2)} + \frac{\exp\left(\frac{-r}{\theta_1}\right)}{(\theta_1+\theta_2)}, \quad 0 < r < \infty. \end{aligned}$$

The cumulative distribution function of $R = \max(X_1, X_2) - \min(X_1, X_2)$ is

$$F_R(r) = 1 - \frac{\left(\theta_1 \exp\left(\frac{-r}{\theta_1}\right) + \theta_2 \exp\left(\frac{-r}{\theta_2}\right)\right)}{\theta_2 + \theta_2}, \quad 0 < r < \infty. \tag{3.13}$$

Therefore, if the random variables X_1 and X_2 are independent, we can use the Markov chain method to compute the ARL since the transition probability matrix can be easily constructed by the use of Equations (3.9), (3.12) and (3.13).

3.4.3 Empirical Distribution Function of Chart Statistic

However, when X_1 and X_2 are not independent, it is quite difficult to come up with an analytical form for the distribution of P_t. To make it possible to verify the ARLs obtained by simulations, we suggest using the empirical cumulative distribution function which is defined by

$$F_n(t) = \frac{\text{number of elements in the sample} \leq t}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{P_i \leq t\}, \tag{3.14}$$

where $\mathbf{1}\{P_i \leq t\}$ is the indicator for fixed t.

To construct an accurate empirical cumulative distribution of P_t in Equation (3.6), we use $n = 1,000,000$ P_t values in Equation (3.14).

3.4.4 Markov Chain Method

Once an empirical cumulative distribution of P_t is constructed, the properties of the EWMA chart of either the MAX-MIN or the CDF Method can be approximated by the Markov chain approach. By discretizing the infinite-state transition probability matrix which is described in Brook and Evans (1972) and Lucas and Saccucci (1990) for normal random variables, the performance of the chart can be measured.

In this study, when it comes to using the Markov chain method, the EWMA control chart of either the MAX-MIN or the CDF Method provides two kinds of control charts: 1) the chart issues a signal when $Z_t < h$ and 2) the chart issues a signal when $Z_t > h$. In other words, in some cases a lower control limit is used while in other cases, an upper control limit is used.

We divide the in-control regions into many sub intervals and denote each sub interval by $S_i, i = 1, 2, \ldots, m$. Throughout this study, we use $m = 8,000$. That is, the number of transient states is 8,000. Moreover, to approximate Z_t, the mid-point of each segment is used. The process Z_t is considered to be in-control when Z_t is within the control limits while it is considered to be out of control when it falls outside the control limits.

For the case where $Z_t > h$, the in-control region is $\{0 \leq Z_t \leq h\}$ (see Figure 3.2(a)). For $j, k = 1, 2, \ldots, m$ and k not equal to 1, the transition probability of Z_t from state j to state k is given by

$$\begin{aligned}
p_{kj} &= P\left[Z_t \in S_k | Z_{t-1} \in S_j\right] = P\left[rP_t + (1-r)Z_{t-1} \in S_k | Z_{t-1} \in S_j\right] \\
&\doteq P\left[rP_t + (1-r)Z_{t-1} \in S_k | Z_{t-1} = s_j\right] \\
&= P\left[2(k-1)\omega - (2j-1)(1-r)\omega < rP_t < 2k\omega - (2j-1)(1-r)\omega\right] \\
&= P\left[\frac{2(k-1)\omega - (2j-1)(1-r)\omega}{r} < P_t < \frac{2k\omega - (2j-1)(1-r)\omega}{r}\right], \qquad (3.15)
\end{aligned}$$

where $s_j = (2j-1)\omega$ represents the midpoint of the jth segment and $\omega = \dfrac{h}{2m}$.

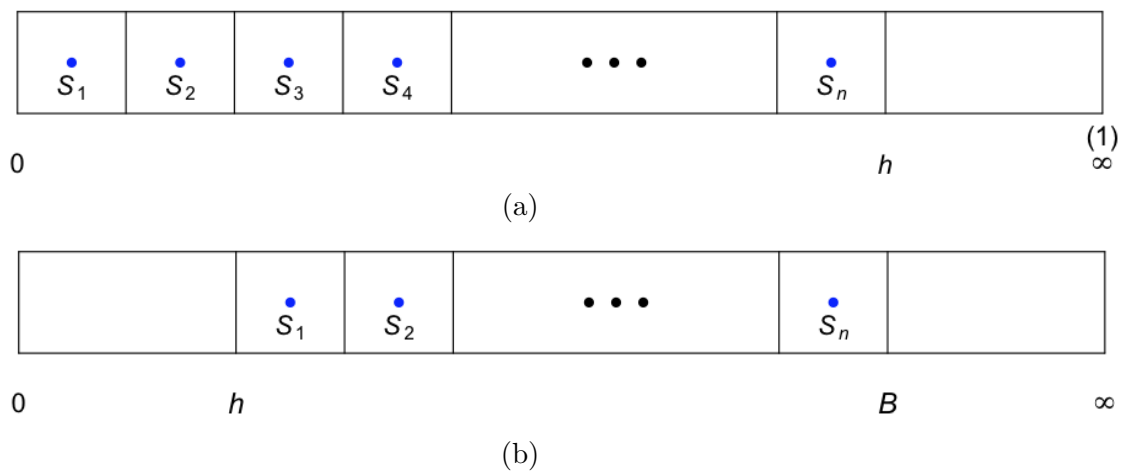


Figure 3.2: An Illustration of Partitioning of the In-Control Region of the EWMA Chart of MAX-MIN and CDF Method (a) $Z_t > h$; (b) $Z_t < h$.

For $k = 1$, we have

$$p_{kj} = P\left[Z_t \in S_k | Z_{t-1} \in S_j\right] \doteq P\left[0 < P_t < \frac{2k\omega - (2j-1)(1-r)\omega}{r}\right]. \tag{3.16}$$

Combining Equations (3.15) and (3.16), the transition probability can be rewritten in the following way:

$$\begin{aligned} p_{kj} &= P\left[Z_t \in S_k | Z_{t-1} \in S_j\right] \doteq P\left[rP_t + (1-r)Z_{t-1} \in S_k | Z_{t-1} = s_j\right] \\ &= P\left[\max\left(0, \frac{2(k-1)\omega - (2j-1)(1-r)\omega}{r}\right) < P_t < \frac{2k\omega - (2j-1)(1-r)\omega}{r}\right], \\ &\quad \text{where } s_j = (2j-1)\omega \text{ represents the midpoint of the } j\text{th segment and } \omega = \frac{h}{2m}. \end{aligned}$$

For the case where $Z_t < h$, the actual in-control region is $\{h \leq Z_t \leq \infty\}$. However, by choosing a large number B which corresponds to 99.9999999999% of the distribution of Z_t, the in-control region can be bounded so that almost every value of Z_t would fall into the in-control region $\{h \leq Z_t \leq B\}$ (see Figure 3.2(b)). Thus, by taking a similar approach, the

transition probability of Z_t from state j to state k is given by

$$\begin{aligned} p_{kj} &= P\left[Z_t \in S_k | Z_{t-1} \in S_j\right] \doteq P\left[rP_t + (1-r)Z_{t-1} \in S_k | Z_{t-1} = s_j\right] \\ &= P\left[2(k-1)\omega + h - (1-r)s_j < rP_t < 2k\omega + h - (1-r)s_j\right] \qquad (3.17) \\ &= P\left[\frac{2(k-1)\omega + h - (1-r)(\omega(2j-1)+h)}{r} < P_t < \frac{2k\omega + h - (1-r)(\omega(2j-1)+h)}{r}\right], \\ &\quad \text{where } \omega = \frac{B-h}{2m} \text{ and } s_j = \omega(2j-1)+h. \end{aligned}$$

All of the transition probabilities in Equations (3.15) - (3.17) can be obtained by using the empirical cumulative distribution function and the transition probability matrix can be constructed, $\mathbf{P} = (p_{kj})_{k=1}^{m}{}_{j=1}^{m}$. The ARL of the control scheme, using the fundamental matrix of the Markov chain, is as follows:

$$\text{ARL} \doteq \mathbf{s}'(\mathbf{I} - \mathbf{P})^{-1}\mathbf{1},$$

where $\mathbf{s}$ denotes the starting vector of the chain and $\mathbf{1}$ denotes the $m \times 1$ vector of 1s.

Tables 3.10 and 3.11 present comparisons of the ARLs obtained by simulations and by the Markov chain method for the detection of D-D shifts of the Exponential-Exponential and the Gamma-Weibull joint distributions. Note that numbers in parentheses are the ARLs obtained by the Markov chain approach and other numbers are simulated ARLs. Control limits in the tables are obtained by simulations and the Markov chain method uses the same control limits to approximate the ARLs. The second row contains the in-control ARLs between two methods while the other rows contain the out-of-control ARLs. The results of Tables 3.10 and 3.11 indicate that there is a good agreement between the ARLs from simulations and the Markov chain approach.

Table 3.10: ARLs Obtained by Simulation and the Markov Chain Method for the D-D Shifts with MAX-MIN Method When Exponential-Exponential Joint Distribution is Used

(δ_1, δ_2)	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	0.5264	04268	0.3065	0.2305	0.1756	0.1331
(1.0 1.0)	ARL_1	200.00	200.00	200.00	200.00	200.00	200.00
	(ARL_1)	(202.01)	(202.53)	(202.08)	(201.18)	(199.43)	(200.23)
(0.9, 0.9)	ARL_1	85.19	95.58	109.99	118.83	125.78	133.26
	(ARL_1)	(85.41)	(96.20)	(110.90)	(118.86)	(126.15)	(131.91)
(0.5, 0.5)	ARL_1	14.38	13.92	14.71	16.51	19.04	22.36
	(ARL_1)	(14.36)	(14.02)	(14.68)	(16.68)	(18.98)	(22.13)
(0.9, 0.6)	ARL_1	28.33	30.85	37.34	44.10	50.80	58.28
	(ARL_1)	(28.30)	(30.93)	(37.41)	(43.83)	(51.03)	(57.65)
(0.5, 0.2)	ARL_1	8.89	7.86	7.01	6.79	6.97	7.46
	(ARL_1)	(8.82)	(7.89)	(7.04)	(6.78)	(6.96)	(7.40)
(1.0, 0.5)	ARL_1	21.30	22.21	25.91	30.48	35.63	41.64
	(ARL_1)	(21.26)	(22.56)	(26.07)	(30.47)	(35.02)	(40.82)
(1.0, 0.2)	ARL_1	9.11	8.09	7.28	7.11	7.36	7.95
	(ARL_1)	(9.07)	(8.11)	(7.28)	(7.08)	(7.38)	(7.93)

Table 3.11: ARLs Obtained by Simulation and the Markov Chain Method for the D-D Shifts with MAX-MIN Method When Gamma-Weibull Joint Distribution is Used

(δ_1, δ_2)	r	0.05	0.10	0.20	0.30	0.40	0.50
	h	2.5559	2.2931	1.9281	1.6539	1.4275	1.2283
(3.54 3.54)	ARL_1	200.00	200.00	200.00	200.00	200.00	200.00
	(ARL_1)	(198.53)	(201.68)	(201.08)	(197.90)	(197.23)	(198.71)
(3.0, 3.0)	ARL_1	22.89	23.37	24.89	26.93	29.04	31.48
	(ARL_1)	(22.94)	(23.15)	(24.71)	(26.67)	(29.36)	(31.59)
(2.0, 2.0)	ARL_1	6.94	6.22	5.51	5.15	4.95	4.87
	(ARL_1)	(6.93)	(6.20)	(5.50)	(5.13)	(4.95)	(4.89)
(1.5, 1.5)	ARL_1	5.21	4.52	3.93	3.42	3.22	3.11
	(ARL_1)	(5.20)	(4.50)	(3.92)	(3.44)	(3.22)	(3.12)
(1.0, 1.0)	ARL_1	4.30	3.61	3.05	2.58	2.44	2.38
	(ARL_1)	(4.29)	(3.59)	(3.07)	(2.59)	(2.45)	(2.39)
(3.0, 2.5)	ARL_1	13.57	13.01	12.66	12.91	13.31	13.84
	(ARL_1)	(13.55)	(13.00)	(12.72)	(12.93)	(13.32)	(13.86)
(4.0, 3.0)	ARL_1	35.05	42.62	57.40	70.31	81.15	90.10
	(ARL_1)	(34.97)	(42.65)	(57.15)	(70.21)	(80.50)	(89.44)
(9.0, 1.0)	ARL_1	1.46	1.27	1.14	1.10	1.08	1.07
	(ARL_1)	(1.42)	(1.27)	(1.14)	(1.10)	(1.08)	(1.07)

3.5 Correlation Study

One remaining aspect of interest is how the control chart driven by either the MAX-MIN or the CDF Method responds to the change in the correlation between random variables

X_1 and X_2. In this section, we examine how the ARL_{opt} values change for each method by trying 4 different dependence parameters representing 0, 0.22, 0.62, and 0.95 correlation between X_1 and X_2. The joint distribution that we use is the Gamma-Weibull distribution.

For illustration, Figure 3.3 plots the ARL_{opt} versus the correlation. Figure 3.3 shows that for the D-D and U-U shifts, smaller ARL_{opt} values are observed, with lower correlation. In other words, smaller correlations lead to better performance of the chart. However, for the U-D shift, a complete opposite result is obtained: smaller ARL_{opt} values are observed when higher correlation is used.

To elicit explanation of these divergent patterns for D-D shifts, box plots of the basic MAX-MIN and CDF statistics P_t values behind the charts were plotted and the graphs are presented in Figure 3.4. In Figure 3.4(a) we notice increasing overlapping among the box plots of MAX-MIN values with increasing correlation, thus making it more difficult to distinguish between the IC and OOC states. As a result, the ARL_1 increases. On the other hand, the Figure 3.4(b) show a greater separation of CDF values between IC and OOC states, thus yielding a quicker detection of the OOC state.

3.6 Discussion

When it comes to the detection of mean changes in bivariate Exponential-type data, the odd phenomenon where the out-of-control ARL becomes bigger than the in-control ARL for the detection of the downward shifts has been recognized by several researchers (Stoumbos and Sullivan 2002, Xie et al. 2011). It is a clear sign that this occurrence is detrimental to the chart detection power and it has to be avoided.

The objectives in this work are 1) to overcome the adverse results and 2) to improve chart performance. To this end, we have proposed two statistics (MAX-MIN, CDF) used

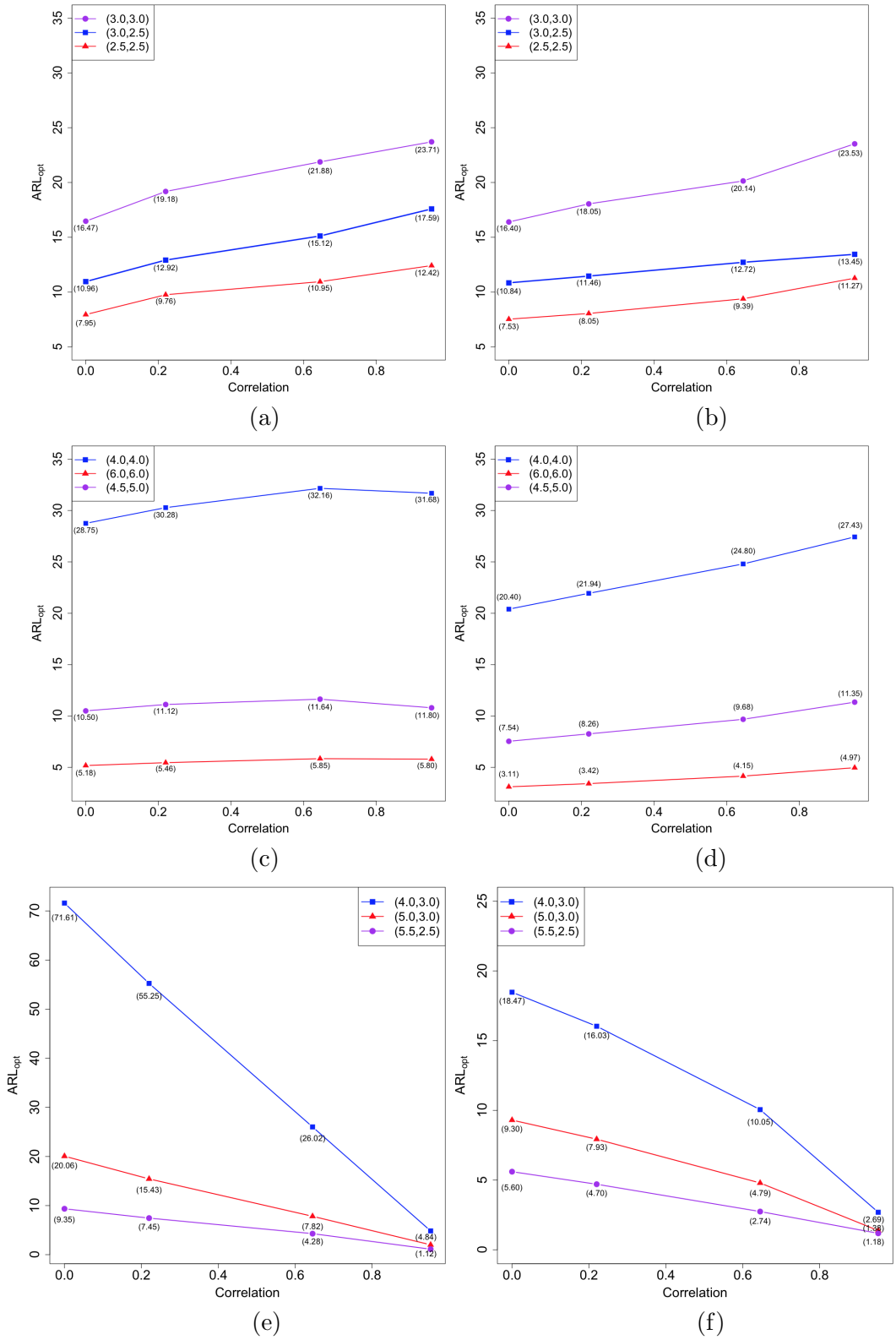
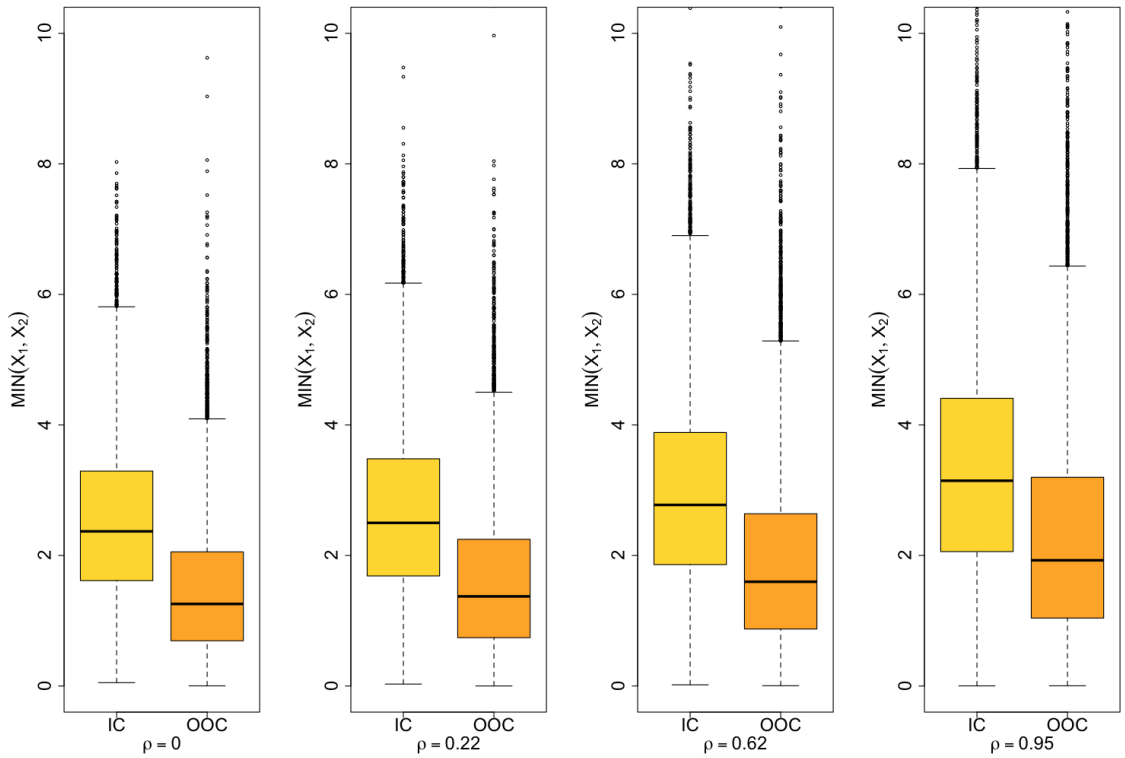
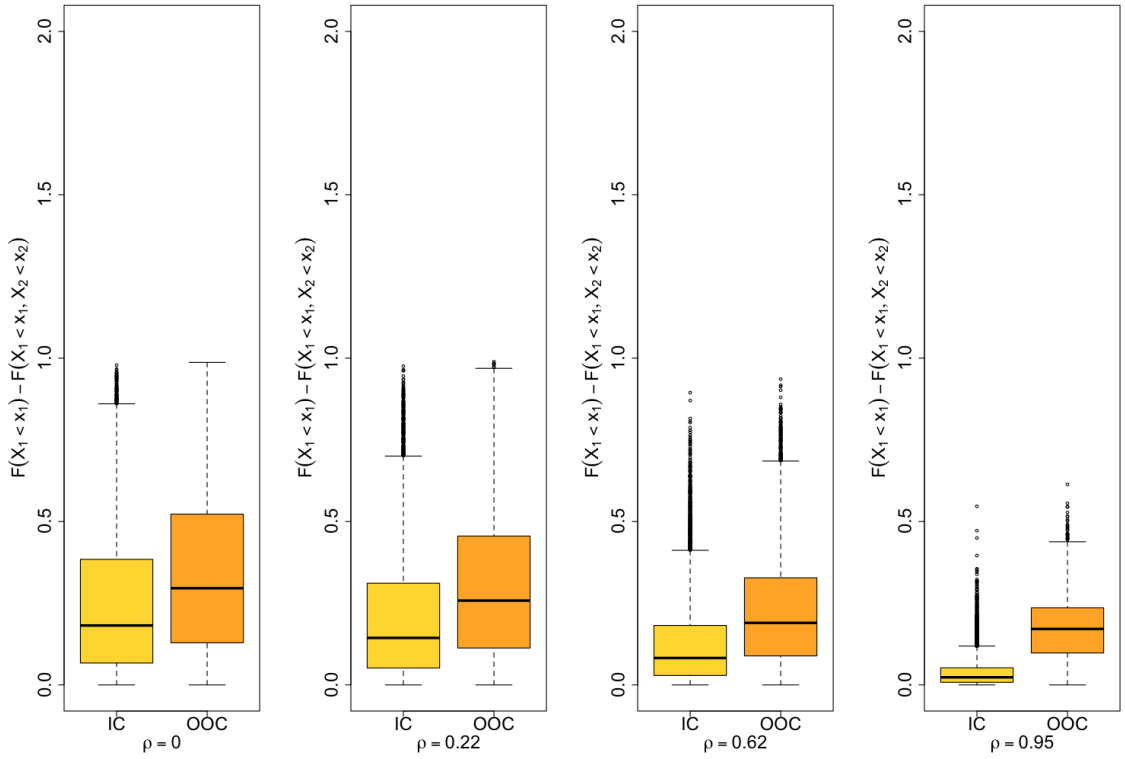


Figure 3.3: Chart Performances Regarding Various Correlations (a) MAX-MIN Method for the D-D Shifts; (b) CDF Method for the D-D Shifts; (c) MAX-MIN Method for the U-U Shifts; (d) CDF Method for the U-U Shifts; (e) MAX-MIN Method for the U-D Shifts; (f) CDF Method for the U-D Shifts.



(a)



(b)

Figure 3.4: Distributional Difference Between IC P_t and OOC P_t (a) MAX-MIN Method for the D-D Shifts; (b) CDF Method for the U-D Shifts.

in conjunction with a univariate EWMA chart. For the generation of skewed data, Xie et al. (2011) studied the bivariate Exponential distribution. Our study extends the methods to the Gamma-Weibull distribution by using the survival Gumbel copula (Gumbel 1960).

Our extensive numerical results show that not only can our proposed methods avoid the strange occurrence but also they can improve chart performance. For example, for the detection of D-D shifts, it is highly recommended that the MAX-MIN Method be used, which is superior to the Reference Method in all directions used. For the detection of U-U shifts, either the MAX-MIN or the CDF Method has an excellent performance in many directions but there are some areas where they do not perform well. However, the same methods with one variable can make our approach more competitive in that area. Finally, for the detection of U-D shifts, the CDF Method improves the chart performance in all directions used.

Apart from the improvement of the chart performance, the proposed methodologies could be easily extended to higher dimensions. Suppose that we have a multivariate control chart with p continuous random variables. As an extreme case, marginal distributions of each variable are all different. By using a copula function (for example, Equation (3.3)), the joint distribution can be easily constructed. The benefit of our proposed methods is that we can reduce p dimensions to a single dimension. In other words, the EWMA charts run by our proposed methods do not suffer from the so-called, "curse of dimensionality" that other control charts might experience in high dimensional space. As a result of dimension reduction, the chart would be relieved of the computational burden arising from a large number of quality indicators.

Chapter 4

Designing Optimal Charts with Prior Information

4.1 Introduction

When the process is out-of-control, the average run length (ARL) is defined as the average number of samples required to detect the out-of-control signal. It has been a common tool to measure the detection power of a control chart. Conventionally, a control chart has been evaluated under the assumption that the mean shift is a fixed and known quantity. However, Ryu et al. (2010) argued that this assumption might not hold in practice because in reality the size of the mean shift δ is rarely known and erroneous specification of the size of the mean shift can lead to poor performance of the control chart. To address this issue, they modelled the mean shift δ with a probability distribution, introduced a new performance measure called the expected weighted run length (EWRL) and discussed the design and optimal performance of various cumulative sum (CUSUM) control charts based on the EWRL.

Another standard practice in control chart calibration has been to work with a time $\tau = 1$ for the process to jump from the in-control state ($\delta = 0$) to an out-of-control state

($\delta \neq 0$). We know, however, that in reality the time τ when the shift in the process mean occurs is actually a random variable. But usually nobody takes into account this fact in chart calibration and evaluation.

In this chapter we extend Ryu et al. (2010)'s approach and also model the transition time as a random variable, denoted by T, whose values follow a discrete distribution over $\{1, 2, 3, \dots\}$. This (δ, τ) becomes a random vector, devoted by (Δ, T), whose probabilistic behaviour is determined by their joint probability density function, $f(\delta, \tau)$. Details are provided later on in the chapter on strategies to elicit $f(\delta, \tau)$.

The primary motivation for our proposed approach is to make both the calibration and evaluation of charts more realistic and relevant for the way processes behave in real life. This will be accomplished by generalizing the EWRL measure of Ryu et al. (2010), as well as the use of other measures we introduce in this chapter, to the context where both δ and τ are subject to uncertainty and select chart settings (e.g. smoothing parameters and control limits) that minimize on the average the time it takes to detect process changes.

In the next section, we discuss various performance measures when both δ and τ are subject to uncertainty and briefly explain about the control charts that we use in this study. Then, we compare the performance of several univariate control charts based on the variants of the EWRL. Finally, the chapter ends with a discussion on the advantages and limitations of the generalized performance measure.

4.2 Performance Measures with Uncertain Shift and Transition Point

Ryu et al. (2010) pointed out that if a control chart is designed based on one specific size of the mean shift, it would perform poorly when the actual size of the mean shift is significantly different from the assumed size. With uncertainty in the mean shift δ, they proposed a new

performance criterion, the EWRL, by assuming that δ follows a probability distribution and discussed various types of EWRL. The EWRL is defined as follows:

$$\begin{aligned} \text{EWRL} &= E[w(\delta)\text{ARL}(\delta)] \\ &= \int_a^b w(\delta)\text{ARL}(\delta)f(\delta)d\delta, \end{aligned} \tag{4.1}$$

where $w(\delta)$ is a weight function associated with δ, ARL(δ) is the average run length for a given departure δ with $\tau = 1$, and $f(\delta)$ is the density function of δ within a range $[a, b]$.

Note that the EWRL can bear different names with different types of weight functions considered. When the weight function is a constant, for example, $w(\delta) = 1$ for all $\delta \in [a, b]$, the EWRL becomes the expected ARL (EARL) with the following form:

$$\begin{aligned} \text{EARL} &= E[\text{ARL}(\delta)] \\ &= \int_a^b \text{ARL}(\delta)f(\delta)d\delta. \end{aligned} \tag{4.2}$$

On the other hand, when $w(\delta) = \dfrac{1}{\text{ARL}_{opt}(\delta)}$, the EWRL is called the expected relative ARL (ERARL) and it has the following form:

$$\text{ERARL} = \int_a^b \frac{\text{ARL}(\delta)}{\text{ARL}_{opt}(\delta)} f(\delta)d\delta. \tag{4.3}$$

In the present work, we suggest that not only could the size of the mean shift be subject to uncertainty just like Ryu et al. (2010) assumed, but also the transition of the mean shift could be random. In other words, we assume that the process mean can depart at any process time (T = τ) and the ARL becomes a function of random variables Δ and T. In general, the in-control ARL (ARL_0) is the average number of observations within the in-control process before a false alarm is issued and the out-of-control ARL (ARL_1) is the average number of observations required to detect a specific mean shift. However, the

proposed ARL is a combination of the in-control ARL and the out-of-control ARL (Huh, 2010). For the computations of ARL, when the shift occurs at time T $= \tau$, we modified the fundamental matrix of the Markov chain and obtain

$$\mathrm{ARL}(\delta,\tau) = \lim_{m_1 \to +\infty} \mathbf{s}' \left\{ \left(\mathbf{I} - \mathbf{P_0}^{\tau-1}\right) (\mathbf{I} - \mathbf{P_0})^{-1} + \mathbf{P_0}^{\tau-1} (\mathbf{I} - \mathbf{P_1})^{-1} \right\} \mathbf{1}, \tag{4.4}$$

where $\mathbf{s}'$ is the starting column vector, $\mathbf{1}$ is a column vector of 1s, $\mathbf{P_0}$ is an on-target transition matrix and $\mathbf{P_1}$ is an off-target transition matrix (see Appendix C). Note that since we use the Markov-chain method, we take $2m_1 + 1$ to be the total number of transient states. Thus, sufficiently large m_1 is recommended for a good approximation. In this study, we use $m_1 = 1500$.

Using the proposed ARL Equation (4.4), the EWRL (see Equation (4.1)) can be generalized as follows. We assume that Δ is a continuous random variable with values on $[a, b]$ and T is a discrete random variable with values $\{c, c+1, c+2, \ldots, d\}$. The new performance measure is then given by

$$\begin{aligned} \mathrm{EWRL} &= E_{(\Delta,T)}[w(\delta,\tau) \cdot \mathrm{ARL}(\delta,\tau)] \\ &= \int_a^b \sum_{\tau=c}^{d} w(\delta,\tau)\mathrm{ARL}(\delta,\tau) f(\delta,\tau) d\delta, \end{aligned} \tag{4.5}$$

where $f(\delta,\tau)$ is the joint probability density function of (Δ, T). Since, as explained before, $\mathrm{ARL}(\delta,\tau)$ has to be computed numerically, there is no analytical expression for EWRL from Equation (4.5). We found it very convenient the use of the usual partition method to approximate Equation (4.5). Specifically, selecting a partition of $[a, b]$, $a = \delta_0 < \delta_1 < \delta_2 < \cdots < \delta_n = b$, and defining $\tilde{\delta}_i = (\delta_{i-1} + \delta_i)/2$, we obtain the approximation

$$\mathrm{EWRL} = \sum_{i=1}^{n} \sum_{\tau=c}^{d} w(\tilde{\delta}_i,\tau)\mathrm{ARL}(\tilde{\delta}_i,\tau) f(\tilde{\delta}_i,\tau) \Delta\delta_i, \tag{4.6}$$

where $\Delta\delta_i = \delta_i - \delta_{i-1}$. We found that a partition of equally-spaced points provided a satisfactory approximation.

With different weight functions, various EWRLs can be introduced. When we consider $w(\delta, \tau) = 1$ for all $\delta \in [a, b]$ and $\tau \in \{c, c+1, c+2, \ldots, d\}$, then the EARL (Equation (4.2)) becomes

$$\begin{aligned} \text{EARL} &= E_{(\Delta,T)}[\text{ARL}(\delta,\tau)] \\ &= \int_a^b \sum_{\tau=c}^{d} \text{ARL}(\delta,\tau) f(\delta,\tau) d\delta \\ &\doteq \sum_{i=1}^{n} \sum_{\tau=c}^{d} \text{ARL}(\tilde{\delta}_i,\tau) f(\tilde{\delta}_i,\tau) \Delta\delta. \end{aligned} \tag{4.7}$$

As a second weight function, consider $w(\delta, \tau) = \dfrac{1}{\text{ARL}_{opt}(\delta,\tau)}$, the ERARL (Equation (4.3)) becomes as follows:

$$\begin{aligned} \text{ERARL} &= E_{(\Delta,T)}\left[\frac{\text{ARL}(\delta,\tau)}{\text{ARL}_{opt}(\delta,\tau)}\right] \\ &= \int_a^b \sum_{\tau=c}^{d} \frac{\text{ARL}(\delta,\tau)}{\text{ARL}_{opt}(\delta,\tau)} f(\delta,\tau) d\delta \\ &\doteq \sum_{i=1}^{n} \sum_{\tau=c}^{d} \frac{\text{ARL}(\tilde{\delta}_i,\tau)}{\text{ARL}_{opt}(\tilde{\delta}_i,\tau)} f(\tilde{\delta}_i,\tau) \Delta\delta, \end{aligned} \tag{4.8}$$

where $\text{ARL}_{opt}(\delta, \tau)$ is obtained from the chart optimally calibrated to detect a mean shift within a range $a \leq \delta \leq b$ and $c \leq \tau \leq d$. In other words, for a particular mean shift δ and a transition point τ, $\text{ARL}_{opt}(\delta, \tau) = \min(\text{ARL}(\delta, \tau))$ when we vary the smoothing parameter $r \in (0, 1]$.

In addition to the EWRL (Equation (4.6)), we propose a couple of new ARL-based performance measures, called the expected effective ARL (EARL^{ef}) and the expected relative effective ARL (ERARL^{ef}), since one might be more interested in the ARL after a transition point of the mean shift. First of all, we define the effective ARL as

$\text{ARL}^{ef} = \text{ARL}(\delta, \tau) - (\tau - 1)$, which is the average number of sampling periods that it takes to detect a change in the mean after it happens. For instance, if $\text{ARL}(\delta, \tau) = 5$ and the mean shift δ takes place at $\tau = 3$, then the $\text{ARL}^{ef} = 5 - (3 - 1) = 3$. It means that on average, three more runs are needed to detect the particular mean shift from the τ value at which the process mean shifted. The EARL^{ef} is given by

$$\begin{aligned} \text{EARL}^{ef} &= \int_a^b \sum_{\tau=c}^{d} \text{ARL}^{ef}(\delta, \tau) f(\delta, \tau) d\delta \\ &\doteq \sum_{i=1}^{n} \sum_{\tau=c}^{d} \text{ARL}^{ef}(\tilde{\delta}_i, \tau) f(\tilde{\delta}_i, \tau) \Delta\delta, \end{aligned} \tag{4.9}$$

where we assume that $w(\delta, \tau) = 1$. Note that when $\tau = 1$, the EARL^{ef} becomes the EARL.

Finally, the last performance measure that we introduce is the expected relative effective ARL (ERARL^{ef}). This is the case where the weight function of δ and τ is the reciprocal of the optimal effective ARL, that is $w(\delta, \tau) = \dfrac{1}{\text{ARL}^{ef}_{opt}(\delta, \tau)}$. The ERARL^{ef} is given by

$$\begin{aligned} \text{ERARL}^{ef} &= \int_a^b \sum_{\tau=c}^{d} \frac{\text{ARL}^{ef}(\delta, \tau)}{\text{ARL}^{ef}_{opt}(\delta, \tau)} f(\delta, \tau) d\delta \\ &\doteq \sum_{i=1}^{n} \sum_{\tau=c}^{d} \frac{\text{ARL}^{ef}(\tilde{\delta}_i, \tau)}{\text{ARL}^{ef}_{opt}(\tilde{\delta}_i, \tau)} f(\tilde{\delta}_i, \tau) \Delta\delta. \end{aligned} \tag{4.10}$$

4.3 Joint Distribution of Δ and T

Eliciting the joint distribution of (Δ, T) required some care. We discuss below two approaches. In the traditional approach to control charts, one relies on Phase I process data to estimate the process parameters (see, e.g. Montgomery 2009, pp 198). These data contain measurements of the quality characteristics of interest from many process runs. The data are usually scrutinized to identify runs suspected to come from out-of-control process

states, the data for these runs are left out and the data from stable runs are used to estimate the process parameters.

One can alternatively use the complete data, including stable and out-of-control runs, and obtain joint values for (δ, τ) from the Phase I process data. Then, using standard statistical methods, one can fit a joint distribution $f(\delta, \tau)$ to the resulting data, for instance, by the method of maximum likelihood. This provides an empirical approach to elicit $f(\delta, \tau)$. If the pairs (δ, τ) exhibit little connection, one can proceed, at least approximately, as though they are independent. In this case, one can fix separate models: $f(\delta)$ for the δ and $f(\tau)$ for the τ data. In this thesis, for the purpose of exploring various joint distributions, we use the specification

$$f_{\Delta,\mathrm{T}}(\delta, \tau) = f_{\Delta|\mathrm{T}}(\delta|\mathrm{T} = \tau)P(\mathrm{T} = \tau), \tag{4.11}$$

where Δ is a continuous random variable and T is a discrete random variable. For computational purposes, first, we assume that T has the negative binomial distribution:

$$\mathrm{T} \sim NB(r, p),$$

where r and p are the distribution parameters. For example, if we use $\tau \in [1 : 300]$ as the range of τ, then the probability mass function of T can be expressed as follows:

$$P_{\mathrm{T}}(\mathrm{T} = \tau) = \frac{\binom{r+\tau-2}{\tau-1}(1-p)^r p^{\tau-1}}{\sum_{\tau=1}^{300}\binom{r+\tau-2}{\tau-1}(1-p)^r p^{\tau-1}} \qquad 1 \leq \tau \leq 300.$$

Regarding the distribution of $\Delta|\mathrm{T}$, we consider the following conditional normal distributions and conditional log-normal distributions:

1. $\Delta|\mathrm{T} = \tau \sim N(\mu = \alpha \pm \beta\tau, \sigma^2)$;

2. $\Delta|\mathrm{T} = \tau \sim \text{Log-}N(\mu = \alpha \pm \beta\tau, \sigma^2)$,

where α, β and σ^2 are distribution parameters. Note that $E[\Delta|\mathrm{T}=\tau]$ and $E[\log(\Delta|\mathrm{T}=\tau)]$ have a linear regression form $(\alpha \pm \beta\tau)$. For example, if the conditional normal distribution with $\delta \in [0.1, 1.5]$ is used, the probability density function is as follows:

$$f_{\Delta|\mathrm{T}=\tau}(\delta) = \frac{\phi(\delta, \mu = \alpha + \beta\tau, \sigma^2)}{\Phi(\delta = 1.5; \mu = \alpha + \beta\tau, \sigma^2) - \Phi(\delta = 0.1; \mu = \alpha + \beta\tau, \sigma^2)}, \quad 0.1 \leq \delta \leq 1.5.$$

For the ERARL and ERARL^{ef}, we assume that Δ and T are independent and they follow uniform distributions for each random variable.

With respect to the ranges of Δ and T, we also use two schemes. For the EARL^{ef} and ERARL^{ef}, we use $\delta \in [0.1, 1.0]$ and $\tau \in [1 : 80]$ while $\delta \in [0.1, 1.5]$ and $\tau \in [1 : 300]$ are used for the EARL and ERARL. Throughout the study, we assume that $\Delta\delta = 0.001$ and the on-target ARL (= ARL$_0$) is 400.

4.4 Control Chart Comparisons

In this section, we present a brief introduction about the charts that we use for performance comparisons. We consider the univariate EWMA, the improved EWMA (IEWMA), the reset EWMA (REWMA) and the one-sided upper CUSUM charts. We assume that observations X_t's are iid following $N(\mu_0, \sigma^2)$. The in-control process mean is μ_0 and the process variance σ^2, does not change over time.

4.4.1 Univariate EWMA Chart by Roberts (1959)

Roberts (1959) introduced the univariate EWMA control chart and its statistic is

$$W_t = rX_t + (1-r)W_{t-1}, \quad t = 1, 2, 3, ..., \tag{4.12}$$

where $W_0 = \mu_0 = 0$ and r is the smoothing parameter that determines the weights of past and current data ($0 < r \leq 1$). The EWMA chart gives an out-of-control signal as soon as $W_t > L\sqrt{\frac{r}{2-r}[1-(1-r)^{2t}]}\sigma$ or $W_t < -L\sqrt{\frac{r}{2-r}[1-(1-r)^{2t}]}\sigma$, where $L > 0$ is chosen to achieve a specified in-control ARL. For the sake of computational simplicity, the asymptotic control limit for the chart $\pm L\sqrt{\frac{r}{2-r}}\sigma$ will be used.

4.4.2 IEWMA Chart by Shu et al. (2007)

The improved one-sided EWMA chart (IEWMA) was proposed by Shu et al. (2007) for quick detection of upward/downward changes in the process mean. The chart statistic is given by

$$X_t^+ = \max[\mu_0, X_t] = \mu_0 + \max[0, X_t - \mu_0]$$

and X_t^+ is applied to the traditional EWMA scheme (see Equation (4.12)). The mean and variance of variable X_t^+ are given by $E[X_t^+] = \mu_0 + \frac{1}{\sqrt{2\pi}}\sigma$, and $\sigma^2_{X_t^+} = \left(\frac{1}{2} - \frac{1}{2\pi}\right)\sigma^2$, respectively. Then, standardizing the statistic X_t^+ by using $Z_t = \frac{X_t^+ - E[X_t^+]}{\sigma_{X_t^+}}$, the IEWMA control chart is defined as

$$S_t = rZ_t + (1-r)S_{t-1}, \quad t = 1, 2, 3, ...,$$

where $S_0 = 0$ and the chart signals if $S_t > h = L\sqrt{\frac{r}{2-r}}$.

4.4.3 REWMA Chart by Crowder and Hamilton (1992)

Crowder and Hamilton (1992) suggested the reset EWMA (REWMA) chart. The REWMA chart is used simply by resetting the traditional EWMA statistic (see Equation (4.12)) to the target whenever it is less than the target. The chart statistic is $Q_t = \max[\mu_0, rX_t + (1-$

$r)Q_{t-1}]$. The chart signals when $Q_t > h = \mu_0 + L\sqrt{\frac{r}{2-r}}\sigma$.

4.4.4 One-Sided Upper CUSUM Chart

The one-sided upper CUSUM chart is defined as

$$C_t^+ = \max[0, C_{t-1}^+ + X_t - k],$$

where k represents the reference value. An alarm is triggered when C_t^+ exceeds the control limit h, which is determined to maintain a given in-control ARL. Note that in general, it is assumed that when the process mean shifts from $\mu_0 = 0$ to μ_1 with the variance 1, the optimal CUSUM chart can be constructed by using $k = \delta/2$. However, in our study, we do not use the assumption since a transition point of the process mean, τ is considered in ARL computations.

4.5 Performance Comparisons

Suppose that the observations of the monitoring process are iid and follow the standard normal distribution with mean 0 and variance 1. In this section, we compare the performance of the univariate EWMA, IEWMA, REWMA and one-sided upper CUSUM charts by considering ERARL, EARL, EARL$_{ef}$ and ERARL$_{ef}$ for given $f(\delta, \tau)$ and $w(\delta, \tau)$. The Markov chain method is employed to obtain the ARL which is a function of δ and τ now, ARL(δ, τ).

4.5.1 ERARL Results

The ERARL is a special case of the EWRL where the weight function $w(\delta, \tau)$ is the reciprocal of the optimal ARL (see Equation (4.8)). Note that we are not allowed to use the reference

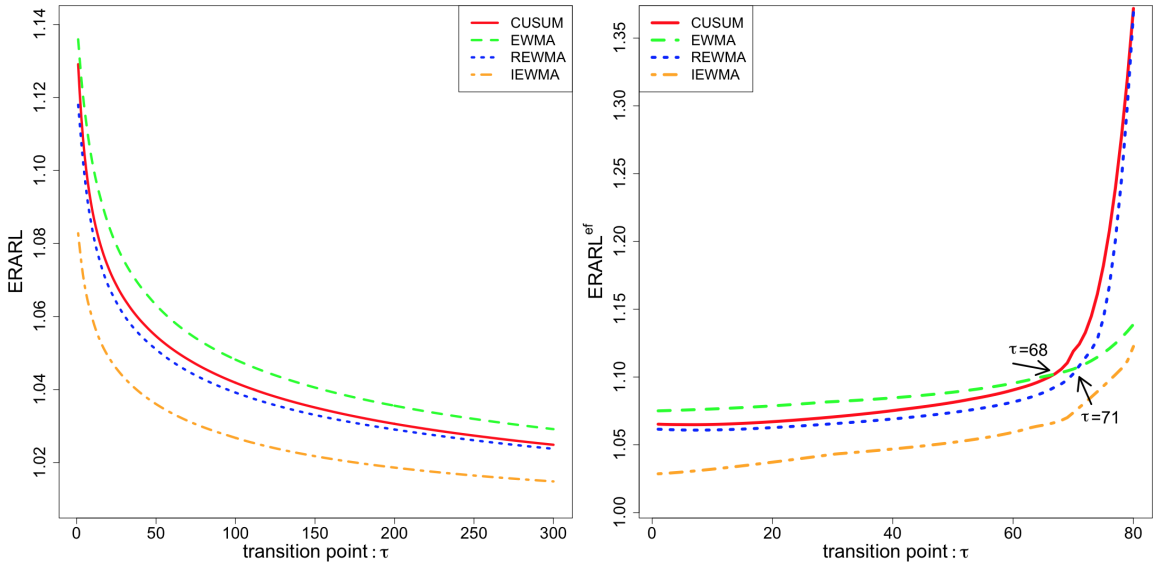
parameter $k = \delta/2$ to optimally construct the CUSUM chart because we assume that not only does the chart performance depend on δ but it is also a function of the transition point τ. In our study, all the optimal ARL values have to be numerically obtained. For the ERARL performance of the control charts mentioned above, we assume that Δ and T are independent and follow uniform distributions. For example, we assume that $\Delta \sim$ continuous $U[0.1, 1.5]$, that is $f(\delta) = \frac{1}{(1.5-0.1)} = \frac{1}{1.4}$ and T $\sim$ discrete $U[1 : 300]$, that is, $P(\tau) = \frac{1}{300}, \tau = 1, 2, \ldots, 300$.

Table 4.1 contains an example of the ERARL$_{opt}$ comparisons in three different ranges of τ ($[1 : 100], [1 : 200]$, and $[1 : 300]$) and it shows that the IEWMA chart outperforms the other charts. Note that in general, when the wider range of τ is used, the smaller ERARL$_{opt}$ is obtained, which means that the performance of control charts improves as a wide range of τ is used. Moreover, given a range of δ which is [0.1, 1.5], Figure 4.1(a) describes how control charts perform as the range of τ increases. In the entire range of τ, the IEWMA chart shows a strong performance while a two-sided control chart EWMA turns out to be the least effective chart among them. Note that with respect to the ERARL$_{opt}$, there is no one chart that outperforms the other chart throughout the range of τ.

Table 4.1: ERARL$_{opt}$ Comparisons for a Variety of τ Ranges. The Range of δ is [0.1, 1.5]

Charts	$\tau = 1$	$\tau : [1 : 100]$	$\tau : [1 : 200]$	$\tau : [1 : 300]$
IEWMA	**1.083**	**1.027**	**1.019**	**1.015**
REWMA	1.118	1.039	1.029	1.024
EWMA	1.136	1.048	1.036	1.029
CUSUM	1.129	1.041	1.031	1.024

- The values in bold indicate the best performance among control charts given a range of τ.



(a) (b)

Figure 4.1: Performances of ERARL and ERARL^{ef} as the Range of τ Increases from $\tau = 1$ to (a) $\tau \in [1 : 300]$; (b) $\tau \in [1 : 80]$, Respectively.

4.5.2 EARL Results

For the estimation of the EARL (Equation (4.7)), we assume that the weight function $w(\delta, \tau) = 1$ and Δ and T are dependent. The joint distribution of Δ and T is given by Equation (4.11). Note that unlike the ERARL computations, we do not consider uniform distributions. The reason is that based on the study by Zhao et al. (2005) and Han et al. (2007), a control chart performing well especially for smaller δ can be a better choice over other control charts. As a result, this scale difference causes the application of uniform distributions to the EARL to become an inappropriate choice.

Table 4.2 describes the performance of the EARL for a variety of joint distributions of Δ and T based upon $\delta \in [0.1, 1.5]$ and $\tau \in [1 : 300]$. The numbers in parentheses in the table are the optimal smoothing parameters for the EWMA-type charts and the reference parameters for the CUSUM control chart. The patterns of the joint distributions of Δ and T are shown as contours in Figures 4.2 - 4.4 and parameter values for each distribution

Table 4.2: EARL Performances for a Variety of Joint Distributions

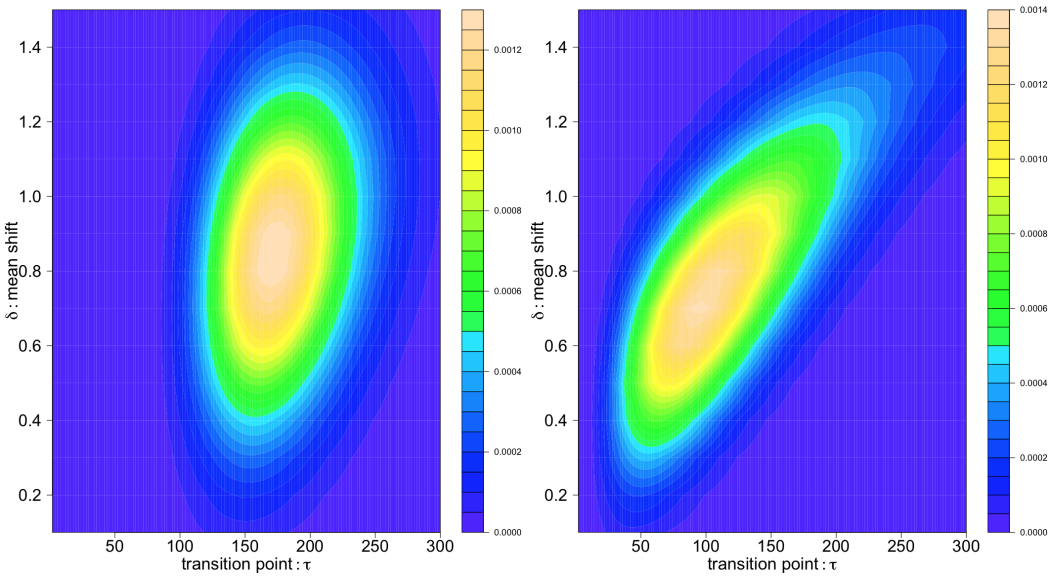
$ARL_0 = 400$				
Distribution Type	IEWMA	REWMA	EWMA	CUSUM
A	**152.25** (0.09)	154.11 (0.13)	155.84 (0.12)	153.59 (0.45)
B	**115.56** (0.07)	117.07 (0.10)	119.39 (0.10)	116.94 (0.36)
C	**118.11** (0.03)	119.05 (0.03)	128.20 (0.03)	118.88 (0.16)
D	**193.04** (0.03)	196.45 (0.04)	204.03 (0.04)	196.22 (0.20)
E	**113.89** (0.07)	115.70 (0.10)	117.85 (0.10)	115.45 (0.37)
F	**131.08** (0.13)	132.85 (0.15)	134.34 (0.13)	132.54 (0.57)

- The values in bold indicate the best performance among control charts.

pattern used are also presented. In a contour plot, the horizontal axis represents the mean transition points (τ) while the vertical axis represents the mean shift size (δ). As Table 4.2 indicates, the IEWMA chart performs better than other charts when distribution types A, B, C, D, E, and F are used (see Figures 4.2 - 4.4). Clearly, the IEWMA chart is the best choice when EARL is used to measure the chart performance. In conclusion, based upon the distributions used in this study, it seems like the IEWMA chart is a better choice.

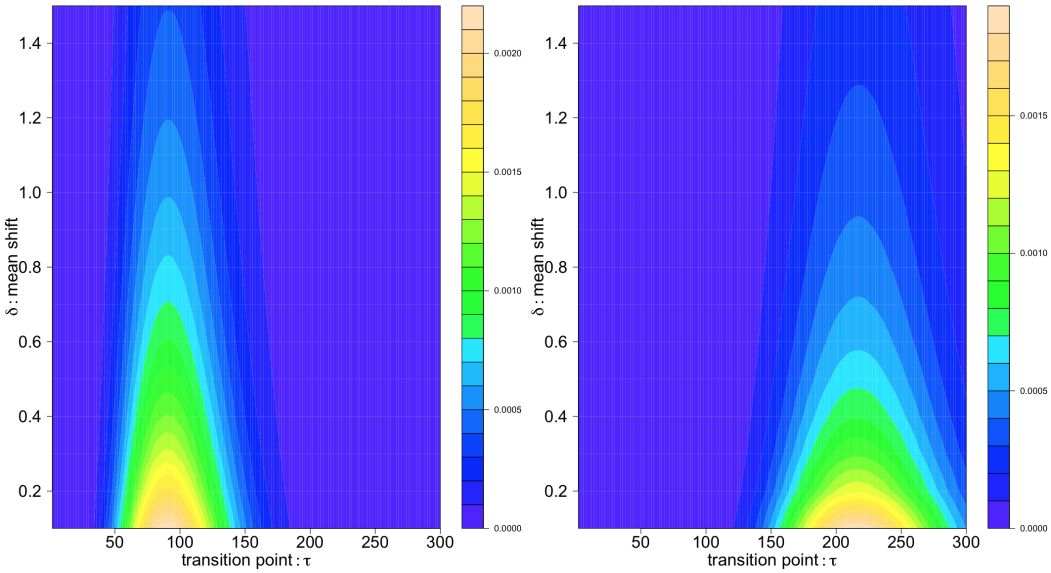
4.5.3 $EARL^{ef}$ Results

One might think that it might be more meaningful to know how many more runs on average are needed to detect a change in the mean after the mean shifts at a certain point of observation (τ), which is now possible with a new performance measure, $EARL^{ef}$ (see Equation (4.9)). Note that in order to get more precise $EARL^{ef}$ estimates, the ranges of δ and τ are reduced to be [0.1, 1.0] and [1 : 80], respectively. Just as with the EARL estimation, we assume that T has the negative binomial distribution and $\Delta|T$ follows either



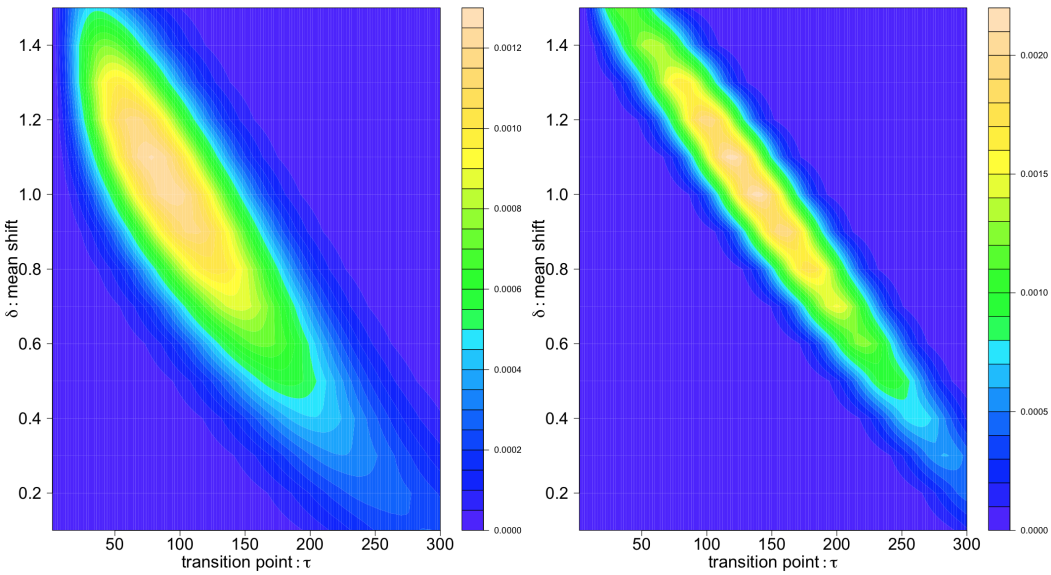
(a) (b)

Figure 4.2: Joint Distributions of Δ and T for EARL (a) Type A: $N(0.55 + \tau/500, 0.3^2), NB(20, 0.1)$; (b) Type B: $N(0.39 + 1.4\tau/370, 0.2^2), NB(4, 0.03)$.



(a) (b)

Figure 4.3: Joint Distributions of Δ and T for EARL (a) Type C: $LN(0.001 + \tau/600, 1.8^2), NB(11, 0.1)$; (b) Type D: $LN(0.005 + \tau/300, 3^2), NB(25, 0.1)$.



(a) (b)

Figure 4.4: Joint Distributions of Δ and T for EARL (a) Type E: $N(1.5 - 1.4\tau/299, 0.2^2), NB(3, 0.022)$; and (b) Type F: $N(1.7 - 1.4\tau/299, 0.1^2), NB(14, 0.023)$.

a normal distribution or a log-normal distribution. The weight function $w(\delta, \tau)$ is assumed to be 1. The patterns of the joint distribution of Δ and T used for the estimation of EARL^{ef} are shown in Figure 4.5. We consider four types of joint distributions of Δ and T (A, B, C and D). Tables 4.3 presents EARL^{ef} values of each control chart for different values of r and k. In this particular example, unlike the results of the EARL where the IEWMA chart was dominant, we have different performance results. The REWMA chart outperforms other charts.

4.5.4 ERARL^{ef} Results

This is another new ARL-based measure, which is the expected relative effective ARL (ERARL^{ef}). The weight function is the reciprocal of its optimal effective ARL (ARL$^{ef}_{opt}$). Just as with the ERARL study, for the estimation of ERARL^{ef} (Equation (4.10)), we assume

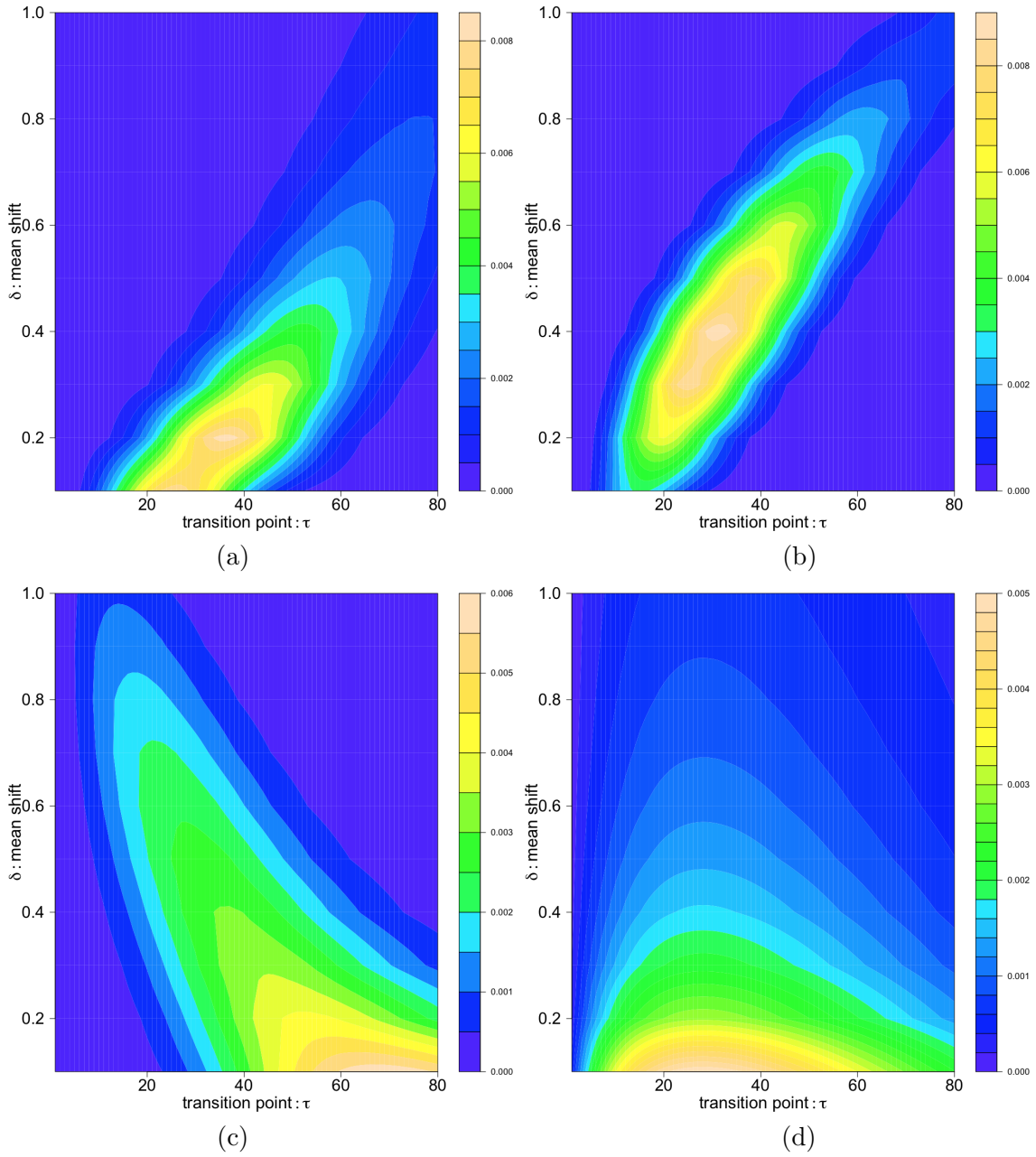


Figure 4.5: Joint Distributions of Δ and T for EARL^{ef} (a)Type A: $LN(-2.40+0.026\tau, 0.38^2)$, $NB(5, 0.09)$; (b)Type B: $N(0.09 + 0.01\tau, 0.11^2), NB(5, 0.12)$; (c)Type C: $N(1.018 - 0.0175\tau, 0.26^2), NB(2, 0.05)$; (d)Type D: $LN(0.001 + 0.02\tau, 10^2), NB(2, 0.035)$.

Table 4.3: EARL^{ef} for a Variety of Joint Distributions

$ARL_0 = 400$				
Distribution Type	IEWMA	REWMA	EWMA	CUSUM
A	41.97 (0.02)	**40.91** (0.02)	54.40 (0.02)	41.34 (0.12)
B	27.61 (0.02)	**26.88** (0.03)	34.65 (0.04)	27.16(0.16)
C	41.64 (0.02)	**40.61** (0.02)	54.26 (0.02)	40.89 (0.11)
D	47.87 (0.02)	**47.06** (0.02)	62.69 (0.02)	47.35 (0.10)

- The values in bold indicate the best performance among control charts.

that Δ and T are independent and each random variable follows uniform distributions so that $\delta \sim$ continuous $U[0.1, 1.0]$ and $\tau \sim$ discrete $U[1:80]$. Figure 4.1(b) shows how control charts react as τ increases. Just as in the study of the ERARL performance (see Figure 4.1(a)), the IEWMA control chart outperforms others throughout the range of τ. However, one distinctive feature is also observed . Unlike in the study of ERARL where control charts behave uniformly throughout the range of τ, control charts do not behave uniformly in this case any more and a couple of interesting upset points ($\tau = 68$ and 71) are observed. Figure 4.1(b) shows that before $\tau = 68$, the EWMA chart is by far the least efficient chart but after $\tau = 71$ it is better than both the CUSUM and REWMA charts.

4.6 Discussion

The ARL is an important basis for measuring the performance of a control chart. Conventionally, the ARL is a function of fixed and known mean shift (δ) and the performance of the control chart has been measured based on the assumption that the mean shift happens at the beginning of the process. Ryu et al. (2010) studied the case where δ is considered as a random variable and proposed a new ARL based measure.

The main objective in this work is to extend Ryu et al. (2010)'s work by taking another random variable T representing the process mean transition point into consideration and propose new performance criteria with the uncertainty in Δ and T. We generalize the EWRL,

proposed by Ryu et al. (2010) and introduce new performance measures such as EARL^{ef}, ERARL^{ef}, EARL and ERARL. The benefit of this approach is that not only can we measure the performance of control charts more precisely when the process starts on-target and then the process mean changes at some time but also predict how many more runs are needed on average to detect a change in the mean after the shift happens.

We tried the new ARL-based criteria on various control charts (IEWMA, REWMA, CUSUM and EWMA charts) under various settings and our numerical results show that the choice of control chart depends on appropriate information of the transition point of the mean shift (τ) and the mean shift (δ). Particularly, when the ERARL^{ef} is used, some interesting phenomenon is observed. For example, until we reach a certain point τ_1, one control chart is better but after the point τ_1, another chart becomes more efficient.

Because the ARL has to be calibrated numerically, all the proposed measures have to be obtained numerically as well. However, the flexibility of choosing the joint distribution of Δ and T, and different weight functions, the new performance measures allow us to easily implement our new approach to various realistic environments. Moreover, in order to use our new approach more effectively, determining an appropriate range of τ becomes important. Especially, the EARL^{ef} and ERARL^{ef} need particular care.

Finally, given δ and τ, the Markov chain method has been used to compute the ARLs. Even though for computational simplicity our method is applied on a univariate control chart in this study, it can easily be extended to higher dimensions. Additionally, since Equation (4.4) is the back-bone of our computations, as long as we can construct a transition probability matrix, our approach can work for any control chart.

Chapter 5

Concluding Remarks

5.1 Summary of Research

Since the introduction of the standard multivariate exponentially weighted moving average (MEWMA) procedure (Lowry et al. 1992), equal smoothing on all quality variables has been conveniently adopted. However, Lowry et al. (1992) warned that this practice should be applied only when there are no a priori reasons to do otherwise. In this thesis, a bivariate exponentially weighted moving average (BEWMA) control statistic with unequal smoothing parameters is introduced with the aim of improving performance over the standard BEWMA chart. Extensive numerical comparisons reveal that the proposed chart enhances the efficiency and flexibility of the control chart in many mean-shift directions. Using the differential smoothing scheme, the ARL depends on both the direction of the process mean shift and the degree of correlation between the quality measurements. The numerical results showed that the proposed chart is at least as good as the standard BEWMA chart and better in many mean shift directions, particularly when the quality variables exhibit low correlations and one wishes to detect subtle mean shifts. Additionally, the new chart shows robustness to correlation misspecification, particularly for low correlations. As a special case where X_1 and X_2 are independent, a new control chart is designed by combining two

independent EWMA control chart and its performance against the bivariate EWMA chart with a single smoothing parameter is investigated.

Recently, Xie et al. (2011) proposed a chart for bivariate Exponential data when the quality measures follow Gumbel's bivariate Exponential distribution (Gumbel 1960). The chart is shown to perform well under the circumstances considered. However, when the process means experience a downward shift (D-D shift), the control charts are shown to break down. In other words, we encounter the strange situation where the out-of-control ARL becomes larger than the in-control ARL. As a result, monitoring small downward mean shifts turns into a tricky job and we are forced to use a small smoothing parameter. Clearly, this is not an ideal approach in certain manufacturing industries such as semiconductor manufacturing industry where monitoring small shifts is highly critical. To address this issue, we have proposed two methods, the MAX-MIN and CDF methods and applied them to the univariate EWMA chart. Our numerical results show that not only do our proposed methods prevent the undesirable behaviour from happening, but they also offer substantial improvement in the ARL over the approach proposed by Xie et al. (2011) in many mean shifts. For the detection of D-D and U-U shifts, the MAX-MIN and the CDF methods are highly recommended, respectively. For U-D (D-U) shifts, both methods are very competitive in many directions. Additionally, we use the empirical cumulative distribution function to construct the distribution of our suggested statistics and verify the simulated ARL values by using the Markov chain method.

Finally, in general, when it comes to designing a control chart, it is assumed that the size of the mean shift is fixed and known. However, Ryu et al. (2010) proposed a new general performance measure, EWRL, by modelling the size of the mean shift with a probability distribution function. We further generalize the measure by introducing a new random variable, T, which is the transition point of the mean shift. Based on that, we propose several ARL-based criteria to measure the chart performance and try them on several univariate control charts. The numerical results show that the distribution of the transition point may

play an important role in choosing the best control chart under the situations considered. Moreover, our proposed criteria help us measure the chart performance when the process starts in-control and then shifts out-of-control later on.

5.2 Future work

Here, some of the future research directions are outlined that could be considered as natural extensions of the work done in this thesis.

5.2.1 The Inertial Properties of Differentially Smoothed Control Charts

Woodall and Mahmoud (2005) proposed a new measure of inertia, the signal resistance, for control charts. Specifically for the EWMA-type charts, whether it is univariate or multivariate, when a small smoothing parameter is used, they recommend using Shewhart-type control limits in conjunction with the charts to reduce the impact of chart inertia. It is reasonable to imply that control charts with differential smoothing will be influenced by the chart inertia, but their research work was conducted based on a single smoothing scheme. The signal resistance for the differential smoothing has to be studied to measure the exact inertia impact on differentially smoothed charts. To some extent, differential smoothing could provide some flexibility to reduce the adverse effect of inertia.

Moreover, in order to counteract the impact of the inertia effect on the differentially-smoothed multivariate EWMA charts, using the fast initial response (FIR) feature could be another approach. However, many studies (Rhoads et al. 1996 and Steiner 1999) about the FIR feature have focused on the univariate EWMA charts. The FIR feature for the charts with differential smoothing scheme could be studied to improve the charts against both large and small shifts without using the Shewhart-type control limits.

5.2.2 The Integral Equation Method under Differential Smoothing

It is generally accepted that there are three ways to obtain the ARL: simulation method, Markov chain method and the integral equation method (Rigdon 1995a, b). The simulation method is the most tedious method for obtaining the ARL, but the advantage of this method is that it can be applied to any situation.

In general, the Markov chain method leads to good results in the ARL computations and this method provides the result very quickly compared to the simulation method. Since the performance of the Markov chain method hinges on the number of transient states, more transient states are required to get a better result. In that sense, when the differential smoothing scheme is employed, a higher level of complexity is expected and this might lead to longer computational time with a large number of transient states. Clearly, under the differential smoothing scheme, a good algorithm will be needed to increase the computational speed.

Another way to compute the ARL analytically is the integral equation method of Rigdon's (1995a, b). But the work so far has been applied to the single smoothing scheme. It is of interest to find out whether the Rigdon's (1995a, b) method can be extended to the differential smoothing scheme.

5.2.3 Differentially Smoothed Control Chart for Nonnormal Multivariate Data

Traditional control charts such as the Shewhart type charts are based on the assumption of normality. At the same time, the single smoothing scheme is conveniently adopted to avoid the complexity which might be induced by differential smoothing. However, skewed data are commonly encountered in manufacturing processes and when several quality characteristics of interest are considered, differential smoothing could be justified. Thus, in that regard, it would be realistic and natural to design a control chart under the differential smoothing

scheme to monitor skewed process data.

5.2.4 Further Application of Copula Functions to Control Charts

In recent years, the study of monitoring nonnormal data has grown. The copula function is a great tool to model a variety of nonnormal joint distributions under certain dependence. Its popularity has been rapidly gaining momentum in the field of econometrics and finance. In our work, we have used the Gumbel copula function (Gumbel 1960) to model and monitor nonnormal data. However, there are other copula functions such as the Clayton copula (Clayton 1978) and the Frank copula (Frank 1979) that could also be applied to control charts. We need to develop guidance on when each copula approach is justified.

5.3 Statement on Computing

In this thesis, the **R** language was used as a main programming language and intensive computations were conducted to obtain the values of ARL. In all three research works, the optimization of a control scheme was performed by going through the following three steps:

(1) program an algorithm in **R** for the computation of ARL

(2) obtain precise control limits by the partition method

(3) find the ARL_{opt} values by varying smoothing parameters for a particular mean shift.

Note that the Step (1) is the most important step and it has to be efficiently programmed and implemented since Steps (2) and (3) are only repetitive operations for which computational times highly depend on Step (1). Moreover, in addition to having fast computing power, high-capacity computers are needed to keep a large number of intermediate results and compare them. In this regard, for the computations of ARL under the differential smoothing scheme and several variants of the EWRL, extra care was taken throughout the computations. For the differential smoothing, due to the additional smoothing parameter, the use of highly nested loop structures was something difficult to avoid and the use of such

control structure in **R** slowed down the computation significantly. For example, in the Intel quad-core i5 (2.7 GHz) platform, it took around 13.5 minutes to get a single value of the ARL in the pure **R** environment. However, when **C** functions were called up from **R** using `.C( )` or `.Call ( )`, the computation time was reduced drastically to around 40 seconds, which is around 20 times faster than the time that the initial **R** program offered. Even though there was a lot of complicated syntax to learn, it was clearly worthwhile to learn how to access **R** functions from **C**. Without the use of **C**, it would have taken a lot longer to complete these tasks.

As to finding the precise control limits, given the $\text{ARL}_0 = 200$ for example, the partition method was used with the precision of the target $\text{ARL}_0 \pm 0.005\%$, which requires that the repetitive operation was submitted to the SHARCNET system and continued until the condition ($199.99 \leq \text{ARL}_0 \leq 200.01$) was met. The partition method brings about combinations of a smoothing parameter and a control limit satisfying a given ARL_0 using the program in Step (1). However, obtaining reliable control limits through simulations was very difficult to maintain the accuracy of $\text{ARL}_0 \pm 0.005\%$. For example, for the second research work where the survival Gumbel copula function was used to monitor bivariate skewed data, a large number of simulations were done to find the control limits with the precision condition of $\text{ARL}_0 \pm 0.5\%$, i.e., ($199 \leq \text{ARL} \leq 201$). Note that for the sake of proper comparisons, the same conditions were applied to other methods such as the standard BEWMA and double BEWMA charts, and obtaining the corresponding control limits with high precision was therefore very time-consuming.

For Step (3), the optimization procedure was the most time-consuming one since every ARL had to be computed over the range of the smoothing parameter ($0 < r \leq 1$) to find the optimal values. To design a control scheme that optimizes the detection performance, this repetitive job was delegated to a SHARCNET cluster such as kraken which has more than 2000 AMD Opteron processors (cores) now. For the differentially-smoothed BEWMA chart, the values of out-of-control ARL needed to be computed given the smoothing parameters

(r_1, r_2), the mean shift and the correlation of quality characteristics. This task was continued until the entire range of smoothing parameters (r_1, r_2) got covered to find the values of ARL_{opt}. For the computations of the variants of EWRL, tasks got more complicated since every optimal value needed to be obtained given the mean shift δ and the transition time τ over the entire range of the smoothing parameter. For example, to obtain the ERARL, first of all, with certain ranges for δ and τ, ARL values were computed and stored in a 400×1399 grid. Note that the reason that we had a 400×1399 grid was because the equal interval $\Delta\delta = 0.001$ and $1 \leq \mathrm{T} \leq 400$ were used. Then, this operation was repeated over the entire range of the smoothing parameter to find the ARL_{opt} for a particular value of δ and τ. The next step was to construct another grid containing a joint probability distribution of Δ and T. This took around more than 4 months to finish the required core computations.

In concluding, I would like to stress that it was a very intensive and time-consuming computational procedure to follow from both software and hardware view points, and it is hard to imagine for me what if the necessary computing powers had not been available. I truly appreciate the use of SHARCNET facility in this regard and also the opportunity to learn **C**.

Appendix A

Corresponding to Chapter 1

A.1 Spherical Distributions

Consider $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3 \ldots$, of standard p dimensional multi-normal observations, with $E\left[\mathbf{X}_t\right] = \mathbf{0}$ and $Var\left[\mathbf{X}_t\right] = \mathbf{I}_p$, $t = 1, 2, \ldots$, and $\mathbf{M}$ is an orthogonal matrix, i.e., ($\mathbf{M}' = \mathbf{M}^{-1}$). By expanding Equation (1.6), we obtain

$$\begin{aligned}
\mathbf{Z}_t &= r\sum_{i=0}^{t-1}(1-r)^i\mathbf{X}_{t-i} + (1-r)^t\mathbf{Z}_0, \quad t = 1, 2, \ldots \\
&= r\sum_{i=0}^{t-1}(1-r)^i\mathbf{X}_{t-i}, \quad \text{since } \mathbf{Z}_0 = \mathbf{0}.
\end{aligned}$$

Then, $\mathbf{Z}_t$ follows a p dimensional multi-normal distribution with $E\left[\mathbf{Z}_t\right] = \mathbf{0}$ and $Var\left[\mathbf{Z}_t\right] = \left[r(1-(1-r)^{2t})/(2-r)\right]\mathbf{I}_p$. When we compute $E\left[\mathbf{M}\mathbf{Z}_t\right]$ and $Var\left[\mathbf{M}\mathbf{Z}_t\right]$, we obtain $E[\mathbf{M}\mathbf{Z}_t] = \mathbf{M}E\left[\mathbf{Z}_t\right] = \mathbf{0}$ and $Var\left[\mathbf{M}\mathbf{Z}_t\right] = \mathbf{M}Var\left[\mathbf{Z}_t\right]\mathbf{M}^{-1} = \left[r(1-(1-r)^{2t})/(2-r)\right]\mathbf{M}\mathbf{M}^{-1} = \left[r(1-(1-r)^{2t})/(2-r)\right]\mathbf{I}_p$. Since the distribution of $\mathbf{M}\mathbf{Z}_t$ is the same as $\mathbf{Z}_t$ for every orthogonal map $\mathbf{M} \in \mathbb{R}^{p\times p}$, $\mathbf{Z}_t$ is a spherical distribution.

Appendix B

Corresponding to Chapter 2

B.1 Derivation of Transition Probabilities

Let us consider

$$\mathbf{W}_t = \mathbf{R}(\mathbf{X}_t - \boldsymbol{\mu}_0) + (\mathbf{I} - \mathbf{R})\mathbf{W}_{t-1}, \quad \text{where } t = 1, 2, ..., \mathbf{W}_0 = \mathbf{0} \text{ and } \boldsymbol{\mu}_0 = \begin{pmatrix} \mu_{10} \\ \mu_{20} \end{pmatrix} \neq \mathbf{0}.$$

The transition probability (the off-target case) of $\mathbf{W}_t$ from state $\mathrm{I}_k \times \mathrm{J}_l$ to state $\mathrm{I}_i \times \mathrm{J}_j$ is given by

$$\begin{aligned}
&P\left(\mathbf{W}_t \in \mathrm{I}_i \times \mathrm{J}_j \middle| \mathbf{W}_{t-1} \in \mathrm{I}_k \times \mathrm{J}_l\right) \\
&= P\left(\begin{pmatrix} r_1(X_1 - \mu_{10}) + (1 - r_1)W_{1,t-1} \\ r_2(X_2 - \mu_{20}) + (1 - r_2)W_{2,t-1} \end{pmatrix} \in \mathrm{I}_i \times \mathrm{J}_j \middle| \begin{pmatrix} W_{1,t-1} \\ W_{2,t-1} \end{pmatrix} \in \mathrm{I}_k \times \mathrm{J}_l\right) \\
&\doteq P\left(\begin{pmatrix} r_1(X_1 - \mu_{10}) + (1 - r_1)\alpha_k \\ r_2(X_2 - \mu_{20}) + (1 - r_2)\beta_l \end{pmatrix} \in \mathrm{I}_i \times \mathrm{J}_j\right),
\end{aligned}$$

where $\alpha_k = -\mathrm{UCL}_1 + (k-0.5)g_1$ and $\beta_l = -\mathrm{UCL}_2 + (l-0.5)g_2$.

$$P\left(\begin{pmatrix} r_1(X_1-\mu_{10}) + (1-r_1)\alpha_k \\ r_2(X_2-\mu_{20}) + (1-r_2)\beta_l \end{pmatrix} \in \mathrm{I}_i \times \mathrm{J}_j\right)$$

$$= P\left(r_1(X_1-\mu_{10}) + (1-r_1)\alpha_k \in \mathrm{I}_i\right)$$

$$\times P\left(r_2(X_2-\mu_{20}) + (1-r_2)\beta_l \in \mathrm{J}_j \middle| r_1(X_1-\mu_{10}) + (1-r_1)\alpha_k \in \mathrm{I}_i\right)$$

$$\doteq P\left(r_1(X_1-\mu_{10}) + (1-r_1)\alpha_k \in \mathrm{I}_i\right)$$

$$\times P\left(r_2(X_2-\mu_{20}) + (1-r_2)\beta_l \in \mathrm{J}_j \middle| r_1(X_1-\mu_{10}) + (1-r_1)\alpha_k = \alpha_i\right)$$

$$= P\left(A_i < r_1(X_1-\mu_{10}) + (1-r_1)\alpha_k \le B_i\right)$$

$$\times P\left(C_j < r_2(X_2-\mu_{20}) + (1-r_2)\beta_l \le D_j \middle| r_1(X_1-\mu_{10}) + (1-r_1)\alpha_k = \alpha_i\right)$$

$$= P\left(\frac{A_i-(1-r_1)\alpha_k}{r_1} + \mu_{10} < X_1 \le \frac{B_i-(1-r_1)\alpha_k}{r_1} + \mu_{10}\right)$$

$$\times P\left(\frac{C_j-(1-r_2)\beta_l}{r_2} + \mu_{20} < X_2 \le \frac{D_j-(1-r_2)\beta_l}{r_2} + \mu_{20} \middle| X_1 = \frac{\alpha_i-(1-r_1)\alpha_k}{r_1} + \mu_{10}\right).$$

When the process is off-target, the conditional distribution of X_2 given X_1 is $X_2|X_1 = a \sim N\Big(\mu_{20} + \Delta x_2 + \frac{\rho\sigma_2}{\sigma_1}(a - \mu_{10} - \Delta x_1), \sigma_2^2(1-\rho^2)\Big)$. Then the transition probability of $\mathbf{W}_t$ (off-target case) from state $\mathrm{I}_k \times \mathrm{J}_l$ to state $\mathrm{I}_i \times \mathrm{J}_j$ is

$$P\left(\mathbf{W}_t \in \mathrm{I}_i \times \mathrm{J}_j \middle| \mathbf{W}_{t-1} \in \mathrm{I}_k \times \mathrm{J}_l\right)$$

$$\doteq \left[\Phi\left(\frac{\frac{B_i-(1-r_1)\alpha_k}{r_1} - \Delta x_1}{\sigma_1}\right) - \Phi\left(\frac{\frac{A_i-(1-r_1)\alpha_k}{r_1} - \Delta x_1}{\sigma_1}\right)\right]$$

$$\times \left[\Phi\left(\frac{\frac{D_j-(1-r_2)\beta_l}{r_2} - \rho\frac{\sigma_2}{\sigma_1}b - \delta^*}{\sigma_2\sqrt{1-\rho^2}} \right) - \Phi\left(\frac{\frac{C_j-(1-r_2)\beta_l}{r_2} - \rho\frac{\sigma_2}{\sigma_1}b - \delta^*}{\sigma_2\sqrt{1-\rho^2}} \right) \right]$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $\delta^* = \Delta x_2 - \rho\frac{\sigma_2}{\sigma_1}\Delta x_1$ and $b = a - \mu_{10} = \dfrac{\alpha_i - (1-r_1)\alpha_k}{r_1} + \mu_{10} - \mu_{10} = \dfrac{\alpha_i - (1-r_1)\alpha_k}{r_1}$.

B.2 Comparison between Simulation and Markov Chain Method

The Markov chain method developed in this chapter for ARL calculation is an approximate method. Its accuracy improves with increasing numbers of states. Recall that the ARL values reported were based on $m_1 = m_2 = 30$, resulting in 1,700 to 3,000 transient states depending on the value of ρ. The question is: how accurate are the resulting ARL values?

ARLs can alternatively be calculated by simulation, although this method can be time consuming, particularly for multivariate charts. To facilitate comparisons with the proposed Markov chain method, the control limit H from the proposed method, obtained by setting $ARL_0 = 200$, was used. The smoothing parameters r_1 and r_2 were set at 0.3 and 0.2, respectively. For each value of H, the on-target and off-target ARLs were computed by simulation for a selection of mean shifts δ_1 and δ_2, as well as by using the proposed Markov chain method. The results are reported in Table B.1 for low, moderate and high levels of correlation, namely $\rho = 0.2$, 0.5 and 0.8. Each case was based on simulating 300,000 run lengths. The simulated ARLs appear in brackets.

A close look at Table B.1 reveals an agreement between the ARLs from the two methods. This applies not only to the different mean shifts considered, but also across all the levels of correlation.

Table B.1: ARL Values Obtained by the Markov Chain Method with $m_1 = m_2 = 30$ and by Simulation

$(r_1, r_2) = (0.30, 0.20)$		$H = 9.8762$		$\rho = 0.2$	$ARL_0 = 200$
$\delta_2 \backslash \delta_1$	-0.4	-0.2	0.0	0.2	0.4
-0.4	36.77 (36.65)	49.47 (49.21)	50.27 (50.15)	38.30 (38.38)	25.31 (25.38)
-0.2	56.60 (56.28)	100.78 (100.23)	118.89 (119.05)	79.46 (79.52)	42.35 (42.42)
0.0	60.63 (60.49)	131.58 (131.35)	200.00 (199.58)	131.58 (131.18)	60.63 (60.51)
0.2	42.35 (42.41)	79.46 (79.58)	118.89 (119.00)	100.78 (100.18)	56.60 (56.21)
0.4	25.31 (25.33)	38.30 (38.23)	50.27 (50.13)	49.47 (49.15)	36.77 (36.53)
$(r_1, r_2) = (0.30, 0.20)$		$H = 9.8693$		$\rho = 0.5$	$ARL_0 = 200$
$\delta_2 \backslash \delta_1$	-0.4	-0.2	0.0	0.2	0.4
-0.4	44.44 (44.15)	51.95 (51.81)	41.30 (41.32)	26.37 (26.48)	16.43 (16.49)
-0.2	62.46 (61.92)	112.10 (111.74)	106.79 (106.76)	57.47 (57.69)	28.59 (28.67)
0.0	50.22 (50.18)	120.21 (119.64)	200.00 (199.74)	120.21 (119.87)	50.22 (50.07)
0.2	28.59 (28.67)	57.47 (57.84)	106.79 (106.15)	112.10 (111.60)	62.46 (62.06)
0.4	16.43 (16.46)	26.37 (26.43)	41.30 (41.44)	51.95 (51.78)	44.44 (44.11)
$(r_1, r_2) = (0.30, 0.20)$		$H = 9.8458$		$\rho = 0.8$	$ARL_0 = 200$
$\delta_2 \backslash \delta_1$	-0.4	-0.2	0.0	0.2	0.4
-0.4	50.85 (52.53)	42.33 (42.31)	22.42 (22.39)	12.41 (12.31)	7.71 (7.67)
-0.2	57.32 (57.98)	120.10 (121.21)	71.45 (71.14)	27.67 (27.62)	12.92 (12.94)
0.0	27.25 (27.42)	85.28 (84.52)	200.00 (200.62)	85.28 (85.82)	27.25 (27.13)
0.2	12.92 (12.70)	27.67 (27.11)	72.59 (73.09)	120.10 (121.00)	57.32 (57.56)
0.4	7.71 (7.63)	12.41 (12.25)	22.78 (22.71)	42.33 (41.99)	50.85 (51.00)

B.3 Comparisons of Number of Transient States between the Single Smoothing and Differential Smoothing Schemes

Since the ARL is approximated by a discrete-space Markov chain, it is clear that the more transient states are factored in a Markov chain approach, the more precise ARL is obtained. Thus, it is important to decide which m_1 and m_2 should be used for the single smoothing and differential smoothing schemes for comparison purposes. Table B.2 shows the number of transient states for various cases, given the on-target ARL is 200. For example, when the differential smoothing scheme is used, the number of transient states varies but the maximum number of transient states is 2933 (when the correlation between X_1 and X_2 is 0). It is observed that the number of transient states gradually start to decrease as the correlation increases. Based on the results obtained in Table B.2, one might choose the setting of $m_1 = 30$ and $m_2 = 60$. However, as Table B.3 shows, the numerical results do not

provide any significant changes in the optimal ARL values even when $m_1 = 30$ and $m_2 = 30$ is used. Consequently, we apply $m_1 = 30$ and $m_2 = 30$ to both the single smoothing and differential smoothing schemes to lessen computational burden.

Table B.2: Number of Transient States

	Differential Smoothing				Single Smoothing		
ρ	0.0	0.25	0.50	0.80	(m_1, m_2)	(30, 30)	(30, 60)
Number of transient states	2933	2825 ∼ 2931	2527 ∼ 2929	1759 ∼ 2909	Number of transient states	1497	2925

B.4 Adding Absorbing States

We show below that transient and absorbing sub-rectangles can be included in an expanded transition probability matrix $\mathbf{P}^*$ using Equations (2.4) and (2.5) for transitions between transient states and 0 for any transition to or from absorbing states. Note that each dimension of $\mathbf{P}^*$ is fixed at $(2m_1 + 1) \times (2m_2 + 1)$. Let $\mathbf{s}^*$ be the $(2m_1 + 1) \times (2m_2 + 1)$ column vector with 1 in the position of the (transient) state containing (0,0) and 0 otherwise. Then,

$$ARL_1 \doteq \mathbf{s}^{*\prime}(\mathbf{I} - \mathbf{P}^*)^{-1}\mathbf{1}. \tag{B.1}$$

The main advantage of computing the ARL in this way is its simplicity in programming. Regarding the on-target ARL (ARL_0), it can be computed by using Equation (2.6) or (B.1) by setting $\Delta x_1 = 0$ and $\Delta x_2 = 0$ in the above.

Consider $\mathbf{P}$, the $n \times n$ transition probability matrix for transient states of some Markov

Table B.3: Comparison of ARL_{opt} Values

$(m_1, m_2)/\ \delta$	0.25	0.50	0.70	1.00	1.50	2.00	2.50	3.00
(30, 30)	62.489	26.682	16.742	9.967	5.429	3.521	2.511	1.878
(30, 60)	62.402	26.667	16.743	9.971	5.433	3.524	2.514	1.881

chain. Let $\mathbf{A} = \mathbf{I} - \mathbf{P}$ and $\mathbf{A}$ and $\mathbf{A}^{-1}$ have explicit forms,

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \quad \text{and} \quad \mathbf{A}^{-1} = \begin{pmatrix} a'_{1,1} & a'_{1,2} & \cdots & a'_{1,n} \\ a'_{2,1} & a'_{2,2} & \cdots & a'_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n,1} & a'_{n,2} & \cdots & a'_{n,n} \end{pmatrix}.$$

Also, let $\mathbf{s}$ be the $n \times 1$ vector with a 1 in the component that corresponds to the starting state (i^{th} position) of the chain and zeros elsewhere, that is, $\mathbf{s} = (0, ..., 0, 1, 0, ..., 0)'$. It is known that the average run length until absorption is

$$ARL = \mathbf{s}'\mathbf{A}^{-1}\mathbf{1} = \sum_{k=1}^{n} a'_{i,k},$$

where $\mathbf{1}$ is the column vector of n ones.

Now augment $\mathbf{P}$ with a dummy column and row of zeros in same position, between rows i and $i+1$ and columns i and $i+1$, say. Denote the new $(n+1) \times (n+1)$ matrix by $\mathbf{P}_{aug}$. Define $\mathbf{B}_{aug} = \mathbf{I}_{aug} - \mathbf{P}_{aug}$, where $\mathbf{I}_{aug}$ is the $(n+1) \times (n+1)$ identity matrix. By direct matrix multiplication one can readily show that

$$\mathbf{B}_{aug}^{-1} = \begin{pmatrix} a'_{1,1} & a'_{1,2} & \cdots & a'_{1,i} & 0 & a'_{1,i+1} & \cdots & a'_{1,n} \\ a'_{2,1} & a'_{2,2} & \cdots & a'_{2,i} & 0 & a'_{2,i+1} & \cdots & a'_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a'_{i,1} & a'_{i,2} & \cdots & a'_{i,i} & 0 & a'_{i,i+1} & \cdots & a'_{i,n} \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a'_{i+1,1} & a'_{i+1,2} & \cdots & a'_{i+1,i} & 0 & a'_{i+1,i+1} & \cdots & a'_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a'_{n,1} & a'_{n,2} & \cdots & a'_{n,i} & 0 & a'_{n,i+1} & \cdots & a'_{n,n} \end{pmatrix}$$

Further, consider the augmented initial vector $\mathbf{s}_{aug}$ obtained from $\mathbf{s}$ by adding a 0 be-

tween entries i and $i+1$. Then, one can verify that

$$\mathbf{s}'_{aug}\mathbf{B}_{aug}^{-1}\mathbf{1}_{aug} = \sum_{k=1}^{n} a'_{i,k} = \mathbf{s}\mathbf{A}^{-1}\mathbf{1},$$

where $\mathbf{1}_{aug}$ is the column vector of $n+1$ ones.

The above formula enables one to calculate average run lengths by using the original transition probability matrix $\mathbf{P}$ or the augmented matrix $\mathbf{P}_{aug}$. Clearly, one can extend the result to adding any number of rows/columns of zeros. We found the formula particularly convenient to handle absorbing states by adding rows/columns of zeros to include any number of such states.

Appendix C

Corresponding to Chapter 4

C.1 Generalized ARL

Let us assume that the process is operating on-target. Following Prabhu and Runger (1996), the Markov chain methods for the MEWMA control chart leads to

$$P(N \geq n) \quad = \quad \lim_{m_1 \to +\infty} \mathbf{s}'\mathbf{P_0}^{n-1}\mathbf{1}, \; n = 1, 2, \ldots \tag{C.1}$$

where N is the run length of the scheme, that is, the number of runs until the false signal is issued for the first time. Here, $\mathbf{s}$ is the starting probability vector. $\mathbf{P_0}$ is the $(2m_1 + 1) \times (2m_1 + 1)$ transition matrix for the Markov chain, and $\mathbf{1}$ denotes a vector of 1s of the dimension $2m_1 + 1$. Then, by using Equation (C.1), the probability mass function for the run length N is

$$\begin{aligned} P(N = n) &= P(N \geq n) - P(N \geq n + 1) \\ &= \lim_{m_1 \to +\infty} \mathbf{s}'\mathbf{P_0}^{n-1}\mathbf{1} - \lim_{m_1 \to +\infty} \mathbf{s}'\mathbf{P_0}^{n}\mathbf{1} \\ &= \lim_{m_1 \to +\infty} \mathbf{s}'\mathbf{P_0}^{n-1}(\mathbf{I} - \mathbf{P_0})\mathbf{1}, \quad n = 1, 2, \ldots, \end{aligned} \tag{C.2}$$

where $\mathbf{P_0}^0 = \mathbf{I}$ and $\mathbf{I}$ is the $(2m_1+1)\times(2m_1+1)$ identity matrix. If the process is off-target at the begining, then Equation (C.2) becomes

$$P(N=n) = \lim_{m_1\to+\infty} \mathbf{s}'\mathbf{P_1}^{n-1}(\mathbf{I}-\mathbf{P_1})\mathbf{1}, \quad n=1,2,\ldots \tag{C.3}$$

Consider the situation where the process goes off-target from $\boldsymbol{\mu}=\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}=\boldsymbol{\mu}_1$ at the time $t=\tau(\geq 1)$ and the change sustains. Thus,

$$\boldsymbol{\mu} = \begin{cases} \boldsymbol{\mu}_0, & t=1,2,\ldots,\tau-1, \\ \boldsymbol{\mu}_1, & t=\tau,\tau+1,\ldots \end{cases}$$

In this case, the transition matrix $\mathbf{P}$ changes as well according to

$$\mathbf{P} = \begin{cases} \mathbf{P_0}, & t=1,2,\ldots,\tau-1, \\ \mathbf{P_1}, & t=\tau,\tau+1,\ldots \end{cases}$$

It means that the process mean stays in-control until $t=\tau-1$ and it shifts out-of-control from $t=\tau$ on. By using Equation (C.3), the new probability mass function, which is a function of δ and τ, can be obtained as follows. For example, for $n=1,2,\ldots,\tau-1$,

$$P(N=n) = \lim_{m_1\to+\infty} \mathbf{s}'\left(\mathbf{P_0}^{n-1}-\mathbf{P_0}^{n}\right)\mathbf{1} = \lim_{m_1\to+\infty} \mathbf{s}'\mathbf{P_0}^{n-1}(\mathbf{I}-\mathbf{P_0})\mathbf{1}.$$

For $n=\tau$,

$$P(N=\tau) = \lim_{m_1\to+\infty} \mathbf{s}'\left(\mathbf{P_0}^{\tau-1}-\mathbf{P_0}^{\tau-1}\mathbf{P_1}\right)\mathbf{1} = \lim_{m_1\to+\infty} \mathbf{s}'\mathbf{P_0}^{\tau-1}(\mathbf{I}-\mathbf{P_1})\mathbf{1}.$$

For $n=\tau+1,\tau+2,\tau+3,\ldots$

$$\begin{aligned} P(N=n) &= \lim_{m_1\to+\infty} \mathbf{s}'\left(\mathbf{P_0}^{\tau-1}\mathbf{P_1}^{(n-1)-\tau+1}-\mathbf{P_0}^{\tau-1}\mathbf{P_1}^{n-\tau+1}\right)\mathbf{1} \\ &= \lim_{m_1\to+\infty} \mathbf{s}'\mathbf{P_0}^{\tau-1}\mathbf{P_1}^{n-\tau}(\mathbf{I}-\mathbf{P_1})\mathbf{1}. \end{aligned}$$

Thus, the probability mass function of run length N is

$$P(N=n) = \lim_{m_1 \to +\infty} \begin{cases} \mathbf{s}'\mathbf{P_0}^{n-1}(\mathbf{I}-\mathbf{P_0})\mathbf{1}, & \text{if } n = 1,\, 2,\, \ldots,\, \tau-1, \\ \mathbf{s}'\mathbf{P_0}^{\tau-1}\mathbf{P_1}^{n-\tau}(\mathbf{I}-\mathbf{P_1})\mathbf{1}, & \text{if } n = \tau,\, \tau+1,\, \ldots. \end{cases} \tag{C.4}$$

Furthermore, by summing up Equation (C.4), Equation (C.5) verifies that the probability mass function is a legitimate one.

$$\begin{aligned} & \sum_{n=1}^{\infty} P(N=n) = \sum_{n=1}^{\tau-1} P(N=n) + \sum_{n=\tau}^{\infty} P(N=n) \\ = & \lim_{m_1 \to +\infty} \mathbf{s}' \left\{ \sum_{n=1}^{\tau-1} \left(\mathbf{P_0}^{n-1}(\mathbf{I}-\mathbf{P_0})\right) + \sum_{n=\tau}^{\infty} \mathbf{P_0}^{\tau-1}\mathbf{P_1}^{n-\tau}(\mathbf{I}-\mathbf{P_1}) \right\} \mathbf{1} \\ = & \lim_{m_1 \to +\infty} \mathbf{s}' \left(\mathbf{I} - \mathbf{P_0}^{\tau-1} + \mathbf{P_0}^{\tau-1}\right) \mathbf{1} \\ = & \ 1. \end{aligned} \tag{C.5}$$

Interestingly, when $\tau = \infty$ or $\tau = 1$, then Equation (C.4) reduces to the following forms:

$$P(N=n) = \lim_{m_1 \to +\infty} \begin{cases} \mathbf{s}'\mathbf{P_0}^{n-1}(\mathbf{I}-\mathbf{P_0})\mathbf{1}, & \text{if } \tau = \infty \\ \mathbf{s}'\mathbf{P_1}^{n-1}(\mathbf{I}-\mathbf{P_1})\mathbf{1}, & \text{if } \tau = 1. \end{cases}$$

This result is consistent with Equations (C.2) and (C.3). Note that Equation (C.4) is a general form of the probability mass function of run length N. Furthermore, using Equation

(C.4), the average run length (ARL_1) can be expressed as follows:

$$\begin{aligned}
\mathrm{ARL}_1 &= E(N) = \sum_{n=1}^{\infty} n f(n) = \sum_{n=1}^{\tau-1} n f(n) + \sum_{n=\tau}^{\infty} n f(n) \\
&= \lim_{m_1 \to +\infty} \sum_{n=1}^{\tau-1} n \mathbf{s}' \mathbf{P_0}^{n-1} (\mathbf{I} - \mathbf{P_0}) \mathbf{1} + \sum_{n=\tau}^{\infty} n \mathbf{s}' \mathbf{P_0}^{\tau-1} \mathbf{P_1}^{n-\tau} (\mathbf{I} - \mathbf{P_1}) \mathbf{1} \\
&= \lim_{m_1 \to +\infty} \mathbf{s}' \left\{ \left(\sum_{n=1}^{\tau-1} n \mathbf{P_0}^{n-1} \right) (\mathbf{I} - \mathbf{P_0}) + \mathbf{P_0}^{\tau-1} \left(\sum_{n=\tau}^{\infty} n \mathbf{P_1}^{n-\tau} \right) (\mathbf{I} - \mathbf{P_1}) \right\} \mathbf{1} \\
&= \lim_{m_1 \to +\infty} \mathbf{s}' \left\{ \left(\mathbf{I} - \mathbf{P_0}^{\tau-1} \right) (\mathbf{I} - \mathbf{P_0})^{-1} + \mathbf{P_0}^{\tau-1} (\mathbf{I} - \mathbf{P_1})^{-1} \right\} \mathbf{1} \\
&\doteq \mathbf{s}' \left\{ \left(\mathbf{I} - \mathbf{P_0}^{\tau-1} \right) (\mathbf{I} - \mathbf{P_0})^{-1} + \mathbf{P_0}^{\tau-1} (\mathbf{I} - \mathbf{P_1})^{-1} \right\} \mathbf{1}. \qquad \text{(C.6)}
\end{aligned}$$

Alternatively, we can derive the generalized ARL_1 (Equation (C.6)) by using the *law of total probability.* $E(N)$ can be written as.

$$E(N) = E(N|N \leq \tau - 1) P(N \leq \tau - 1) + E(N|N \geq \tau) P(N \geq \tau).$$

From Equation (C.1), we know that

$$\begin{aligned}
P(N \geq \tau) &= \lim_{m_1 \to +\infty} \mathbf{s}' \mathbf{P_0}^{\tau-1} \mathbf{1} \\
P(N \leq \tau - 1) &= \lim_{m_1 \to +\infty} \left(1 - \mathbf{s}' \mathbf{P_0}^{\tau-1} \mathbf{1} \right).
\end{aligned}$$

The conditional probality of N, given $N \leq \tau - 1$, is

$$f(n|N \leq \tau - 1) = \frac{f(n)}{P(N \leq \tau - 1)} = \lim_{m_1 \to +\infty} \frac{\mathbf{s}' \mathbf{P_0}^{n-1} (\mathbf{I} - \mathbf{P_0}) \mathbf{1}}{1 - \mathbf{s}' \mathbf{P_0}^{\tau-1} \mathbf{1}}, \quad n = 1, 2, \ldots, \tau - 1.$$

The conditional expectation of N, given $N \leq \tau - 1$, is

$$\begin{aligned}
E(N|N \leq \tau - 1) &= \sum_{n=1}^{\tau-1} n f(n|N \leq \tau - 1) = \lim_{m_1 \to +\infty} \sum_{n=1}^{\tau-1} n \frac{\mathbf{s}' \mathbf{P_0}^{n-1} (\mathbf{I} - \mathbf{P_0}) \mathbf{1}}{1 - \mathbf{s}' \mathbf{P_0}^{\tau-1} \mathbf{1}} \qquad \text{(C.7)} \\
&= \lim_{m_1 \to +\infty} \frac{1}{1 - \mathbf{s}' \boldsymbol{P_0}^{\tau-1} \mathbf{1}} \mathbf{s}' \left\{ (\mathbf{I} - \mathbf{P_0}^{\tau-1}) (\mathbf{I} - \mathbf{P_0})^{-1} - (\tau - 1) \mathbf{P_0}^{\tau-1} \right\} \mathbf{1}.
\end{aligned}$$

In the same way, the conditional expectation of N, given $N \geq \tau$, is

$$\begin{aligned} E(N|N \geq \tau) &= \sum_{n=\tau}^{\infty} n f(n|N \geq \tau) = \lim_{m_1 \to +\infty} \sum_{n=\tau}^{\infty} n \frac{\mathbf{s}'\mathbf{P_0}^{\tau-1}\mathbf{P_1}^{n-\tau}(\mathbf{I}-\mathbf{P_1})\mathbf{1}}{\mathbf{s}'\mathbf{P_0}^{\tau-1}\mathbf{1}} \\ &= \lim_{m_1 \to +\infty} \frac{1}{\mathbf{s}'\mathbf{P_0}^{\tau-1}\mathbf{1}} \mathbf{s}' \left\{ \mathbf{P_0}^{\tau-1}(\mathbf{I}-\mathbf{P_1})^{-1} + (\tau-1)\mathbf{P_0}^{\tau-1} \right\} \mathbf{1}. \quad \text{(C.8)} \end{aligned}$$

Therefore, by combining Equations (C.7) and (C.8), we obtain the same result as Equation (C.6):

$$\begin{aligned} \mathrm{ARL}_1 &= E(N) = E(N|N \leq \tau-1)P(N \leq \tau-1) + E(N|N \geq \tau)P(N \geq \tau) \\ &= \lim_{m_1 \to +\infty} \mathbf{s}' \left\{ \left(\mathbf{I}-\mathbf{P_0}^{\tau-1}\right)(\mathbf{I}-\mathbf{P_0})^{-1} + \mathbf{P_0}^{\tau-1}(\mathbf{I}-\mathbf{P_1})^{-1} \right\} \mathbf{1}. \\ &\doteq \mathbf{s}' \left\{ \left(\mathbf{I}-\mathbf{P_0}^{\tau-1}\right)(\mathbf{I}-\mathbf{P_0})^{-1} + \mathbf{P_0}^{\tau-1}(\mathbf{I}-\mathbf{P_1})^{-1} \right\} \mathbf{1}. \end{aligned}$$

Appendix D

Programming (R- code)

D.1 Corresponding to Chapter 2

D.1.1 Markov Chain Algorithm for Differential Smoothing

Note that the following code for the computation of ARL is constructed based on that

$$\mathbf{W}_t = \mathbf{R}\mathbf{X}_t + (\mathbf{I} - \mathbf{R})\mathbf{W}_{t-1}, \quad t = 1, 2, 3, ...,$$

where $\mathbf{W}_0 = \mathbf{0}$ and $\boldsymbol{\mu}_0 = \mathbf{0}$.

```
delta1<-(-0.24494766) # off- target mean of X1

delta2<-(-0.07916700) # off- target mean of X2

mu1<-0 # on-target mean of  X1

mu2<-0 # on-target mean of  X2

sigma1<-1 # the variance of X1

sigma2<-1 # the variance of X2

rho<-(0.5) # the correlation of X1 and X2

sigma12<-rho*sqrt(sigma1)*sqrt(sigma2) # the covariance of X1 and X2
```

```
#Define the smoothing parameters
r1<-0.03
r2<-0.01

#Define the covariance matrix of W
w_sigma1<-r1/(2-r1)*sigma1
w_sigma12<-r1*r2/(r1+r2-r1*r2)*sigma12
w_sigma2<-r2/(2-r2)*sigma2
W<-matrix(c(w_sigma1, w_sigma12, w_sigma12, w_sigma2),nrow=2,ncol=2, byrow=TRUE)
IW<-solve(W)

#Define the number of transient states
m1<-30
m2<-30
H<- 5.316563 # the control limit given the on-target ARL
UCL1<-sqrt(IW[2,2]*H/(IW[1,1]*IW[2,2]-IW[1,2]^2))
UCL2<-sqrt(IW[1,1]*H/(IW[1,1]*IW[2,2]-IW[1,2]^2))
g1<-2*UCL1/(2*m1+1) #the interval of each state along W1 axis
g2<-2*UCL2/(2*m2+1) #the interval of each state along W2 axis

# Transitional matrix
P<-matrix(data=NA,nrow=(2*m1+1)*(2*m2+1),ncol=(2*m1+1)*(2*m2+1))
M<-matrix(data=NA,nrow=(2*m1+1),ncol=(2*m1+1))
V<-matrix(data=NA,nrow=(2*m2+1),ncol=(2*m2+1))

# Range of W1 and W2
range1<-(2*m1+1)
```

```
range2<-(2*m2+1)

# Computation of transition probability
for (k in 1:range1){
ak<-(-UCL1 + (k-0.5)*g1)
for (i in 1: range1){
up_w1<-(-UCL1 + i*g1 - (1-r1)*ak)/r1
down_w1<-(-UCL1 + (i-1)*g1 - (1-r1)*ak)/r1
M[k,i]<-pnorm(up_w1- delta1, mean=0, sd=1) - pnorm(down_w1- delta1, mean=0, sd=1)
ai<-(-UCL1 + (i-0.5)*g1)
for (l in 1: range2){
bl<-(-UCL2 + (l-0.5)*g2)
for (j in 1: range2){
up_w2<-(-UCL2 + j*g2 - (1-r2)*bl)/r2
down_w2<-(-UCL2 + (j-1)*g2 - (1-r2)*bl)/r2
condi_mu<-mu2 + sigma12/sigma1*((ai - (1-r1)*ak)/r1 -delta1 - mu1)
condi_var<-sigma2 - sigma12^2/sigma1
V[l,j]<-pnorm((up_w2-condi_mu-delta2)/sqrt(condi_var),mean=0,sd=1) -
pnorm((down_w2-condi_mu-delta2)/sqrt(condi_var),mean=0,sd=1)
temp<-M[k,i]*V[l,j]
P[(k-1)*(2*m2 + 1)+l,(i-1)*(2*m2 + 1)+j]<-temp
}
}
}
}
P1<-P
counter<-1
```

```
# Removal of absorbing states
for (alpha in 1:range1){
for (beta in 1:range2){
if (IW[1,1]*(alpha-(m1+1))^2*(g1)^2 + 2*IW[1,2]*(alpha-(m1+1))
*(beta-(m2+1))*g1*g2 + IW[2,2]*(beta-(m2+1))^2*(g2)^2 >= H){
P1[,counter]<-0
P1[counter,]<-0
}
counter<-counter+1
}
}
z1<-(2*m1+1)*(2*m2+1)
n1<-c(z1)
I <- matrix(0,nrow=n1,ncol=n1)
I[row(I)==col(I)]<-1 # the identity matrix
one<- matrix(1,nrow=z1,ncol=1) # the 1 vector
temp<-solve(I-P1)
S<- matrix(0,nrow=z1,ncol=1)
start<-m1*(2*m2+1)+m1+1 # the location of the initial transient state
S[start,1]<-1 # the initial vector
# Computation of the ARL by using the fundamental matrix
ARL<-t(S)%*% temp%*%one
ARL
```

D.1.2 Markov Chain Algorithm for Single Smoothing

```
delta<- 0 # off- target mean of X
h<- 4.938465 # the control limit given the on-target ARL
r<- 0.016 # the smoothing parameter
UCL<-sqrt(h*r/(2-r))
#Define the number of transient states
m1<-30
m2<-30
g1<-2*UCL/(2*m1+1)
g2<-2*UCL/(2*m2+1)
p<-2 # the number of quality characteristics of interest
H<-matrix(data=NA,nrow=2*m1+1,ncol=2*m1+1)

z<-(2*m1+1)*(m2+1)
n1<-c(z)
I <- matrix(0,nrow=n1,ncol=n1) # the identity matrix
I[row(I)==col(I)]<-1
one<- matrix(1,nrow=z,ncol=1) # the 1 vector

# Computation of transition probability
range1<-2*m1+1
for (i in 1:range1){
c_i<- -UCL+(i-0.5)*g1
for (j in 1:range1){
up<-(-UCL+j*g1-(1-r)*c_i)/r-delta
```

```
down<-(-UCL+(j-1)*g1-(1-r)*c_i)/r-delta
H[i,j]<-pnorm(up,mean=0,sd=1)-pnorm(down,mean=0,sd=1)
}
}
range2<-m2+1
V<-matrix(data=NA,nrow=range2,ncol=range2)
for (i in 0:m2){
c<-((1-r)*i*g2/r)^2
for (j in 0:m2){
if (j==0) {
V[i+1,1]<-pchisq((0.5*(g2)/r)^2, df=p-1,ncp=c)
}
else {
up<-((j+0.5)*g2/r)^2
down<-((j-0.5)*g2/r)^2
V[i+1,j+1]<-pchisq(up,df=p-1,ncp=c)-pchisq(down,df=p-1,ncp=c)
}
}
}
E<- kronecker(H,V) # computation of the kronecker product

counter<-1
# Removal of absorbing states
for (alpha in 1:range1){
for (beta in 0:m2){
if ((alpha-(m1+1))^2*g1^2+(beta*g2)^2 >= UCL^2){
E[,counter]<-0
```

```
E[counter,]<-0
}
counter<-counter+1
}
}
temp<-solve(I-E)
S<- matrix(0,nrow=z,ncol=1)
start<-m1*(m2+1)+1 # the location of the initial transient state
S[start,1]<-1 # the initial vector
# Computation of the ARL by using the fundamental matrix
ARL<-t(S)%*% temp%*%one
ARL
```

D.2 Corresponding to Chapter 3

D.2.1 Generating Samples from the Bivariate Gumbel Copula

```
N<-1000000  # number of observations
COP<-14 # copula family
dp<-2 # dependence parameter

simdata<-BiCopSim(N, COP, dp) #Data generation from a copula

X1<-NULL
X2<-NULL
Rate1<-1
```

```
Rate2<-1

# Quantile function returns vector of quantiles.
for (i in 1:N){
X1[i]<-qexp(simdata[i,1],rate=Rate1)
X2[i]<-qexp(simdata[i,2],rate=Rate2)
}
```

D.2.2 Markov Chain Algorithm in the Case When $Z_t < h$

```
M<-8000 # number of transient states
range<-M

# Computation of  transition probability with the empirical CDF
G<-matrix(data=NA,nrow=range, ncol=range)
for(j in 1:range){
lambda<-0.02 # the smoothing parameter
UCL<-0.4131 # the control limit
for (k in 1:range){
delta<-UCL/(2*M)
up<-max(0, (2*k*delta-(1-lambda)*((2*j-1)*delta))/lambda-shift)
down<-max(0,(2*(k-1)*delta-(1-lambda)*((2*j-1)*delta))/lambda-shift)
G[j,k]<-Fn(up)- Fn(down)
}
}
```

```
one<- matrix(1,nrow=range,ncol=1)
I2 <- matrix(0,nrow=range,ncol=range) # the identity matrix
I2[row(I2)==col(I2)]<-1

ORG<-0.3754859 # this value is obtained from simulations; the mean of Pt
ii<-1
VA<-delta*2 *ii
# Computation of the initial transient state
while (VA < ORG){
ii<-ii+1
VA<-delta*2*ii
 }
start<-ii

S<- matrix(0,nrow=range,ncol=1)
S[start,1]<-1 # the initial vector
temp<-solve(I2-G)
# Computation of the ARL by using the fundamental matrix
ARL<-t(S)%*%temp%*%one
start
ARL
```

D.2.3 Markov Chain Algorithm in the Case When $Z_t > h$

```
M<-8000 # number of transient states
range<-M
```

```
# Computation of  transition probability with the empirical CDF
G<-matrix(data=NA,nrow=range, ncol=range)
for(j in 1:range){
lambda<-0.02 # the smoothing parameter
UCL<-0.14275 # the control limit
for (k in 1:range){
delta<-UCL/(2*M)
up<-max(0, (2*k*delta-(1-lambda)*((2*j-1)*delta))/lambda-shift)
down<-max(0,(2*(k-1)*delta-(1-lambda)*((2*j-1)*delta))/lambda-shift)
G[j,k]<-Fn(up)- Fn(down)
}
}

one<- matrix(1,nrow=range,ncol=1)
I2 <- matrix(0,nrow=range,ncol=range) # the identity matrix
I2[row(I2)==col(I2)]<-1

ORG<-0.1247858 # this value is obtained from simulations; the mean of Pt
ii<-1
VA<-delta*2 *ii
# Computation of the initial transient state
while (VA < ORG){
ii<-ii+1
VA<-delta*2*ii
 }
start<-ii
```

```
S<- matrix(0,nrow=range,ncol=1)
S[start,1]<-1 # the initial vector
temp<-solve(I2-G)
# Computation of the ARL by using the fundamental matrix
ARL<-t(S)%*%temp%*%one
start
ARL
```

D.3 Corresponding to Chapter 4

D.3.1 Markov Chain Algorithm for the IEWMA Control Chart

```
mu0<-0 # the mean is fixed at 0 without loss of generality
sigma2<-1 # the variance is 1
sigma<-sqrt(sigma2)
L<- 2.632105 # control limit
r<-0.07 # smoothing parameter

sigma_plus<-sqrt(0.5-0.5/pi)*sigma
h<-L*sqrt(r/(2-r))

mu1<- 0
m<- 3001 # number of transient states
y<-0
x<-0
w<-(h+1/sqrt(pi-1))/m
```

```
for (i in 1:m){
if (-1/sqrt(pi-1) + (i-1)*w < y & y < -1/sqrt(pi-1) + i*w ){x<-i}
}
G<-matrix(data=NA,nrow=m,ncol=m)
w<-(h+1/sqrt(pi-1))/m

# Computation of  transition probability

for (i in 1:m){
for (j in 1:m){
a1<-(j-1-(1-r)*(i-0.5))*w*sigma_plus/r+mu0
a2<-(j-(1-r)*(i-0.5))*w*sigma_plus/r+mu0
if (a2 < 0)
{G[i,j]<- 0}
else if (a2 >= 0 & a1 < 0)
{G[i,j]<- pnorm((a2-mu1)/sigma,mean=0, sd=1)
}
else if (a2 >= 0 & a1 >=0)
{G[i,j]<- pnorm((a2-mu1)/sigma,mean=0, sd=1)-pnorm((a1-mu1)/sigma,mean=0, sd=1)
}
}
}

One<- matrix(1,nrow=m,ncol=1) # one vector
I2 <- matrix(0,nrow=m,ncol=m)
I2[row(I2)==col(I2)]<-1 # the identity matrix
```

```
S<- matrix(0,nrow=m,ncol=1)
start<-x
S[start,1]<-1
temp<-solve(I2-G)

# Computation of the ARL by using the fundamental matrix

ARL<-t(S)%*%temp%*%One
ARL
```

D.3.2 Markov Chain Algorithm for the One-Sided CUSUM Control Chart

```
sigma<-1 # the variance is 1
m<-3001 # the number of transient states
delta<- 0 # deviation
mu<-0 # the mean is fixed at 0 without loss of generality

ARL_Calculator<-function(r1, b){
K<-r1 # reference parameter value
H<-b # control limit

# Computation of  transition probability

G<-matrix(data=NA,nrow=m,ncol=m)
w<-2*H/(2*m-1)
for (i in 1:m){
```

```
for (j in 1:m){
if (j == 1)
{G[i,j]<-pnorm(K-(i-1)*w+0.5*w-delta-mu,mean=0, sd=1)}
else
{G[i,j]<-pnorm((j-i)*w+0.5*w+K-delta-mu,mean=0,sd=1)
-pnorm((j-i)*w-0.5*w+K-delta-mu,mean=0, sd=1)
}
}
}

One<- matrix(1,nrow=m,ncol=1) # one vector
I2 <- matrix(0,nrow=m,ncol=m) # the identity matrix
I2[row(I2)==col(I2)]<-1

S<- matrix(0,nrow=m,ncol=1)
start<-1
S[start,1]<-1
temp<-solve(I2-G)

# Computation of the ARL by using the fundamental matrix

ARL<-t(S)%*%temp%*%One
ARL
}
```

D.3.3 Markov Chain Algorithm for the REWMA Control Chart

```
sigma<-1 # the variance is 1
r<-0.2 # the smoothing parameter
L<- 2.79138 # control limit
UCL<-L*sqrt(r/(2-r))
delta<- 0 # the off-target mean deviation
m<- 3001 # the number of transient states
C<- 0

# Computation of  transition probability

G<-matrix(data=NA,nrow=m,ncol=m)
w<-(UCL+C)/m
for (i in 1:m){
for (j in 1:m){
a1<- -C+(j-1-(1-r)*(i-0.5))/r*w
a2<- -C+(j-(1-r)*(i-0.5))/r*w
if (j == 1)
{G[i,j]<-pnorm(a2-delta,mean=0, sd=1)}
else
{G[i,j]<-pnorm(a2-delta,mean=0, sd=1)-pnorm(a1-delta,mean=0, sd=1)
}
}
}
```

```
One<- matrix(1,nrow=m,ncol=1) # one vector

I2 <- matrix(0,nrow=m,ncol=m) # the identity matrix

I2[row(I2)==col(I2)]<-1


S<- matrix(0,nrow=m,ncol=1)

start<-1

S[start,1]<-1

temp<-solve(I2-G)

# Computation of the ARL by using the fundamental matrix

ARL<-t(S)%*%temp%*%One

ARL
```

Bibliography

[1] Alkahtani, S., and Schaffer, J. (2012). "A Double Multivariate Exponentially Weighted Moving Average (dMEWMA) Control Chart for a Process Location Monitoring", *Communication in Statistics - Simulation and Computation* **41**, pp. 238-252.

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

[3] Clayton, D. G. (1978). "A Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence", *Biometrika* **65**, pp. 141-151.

[4] Crosier, R. B. (1988). "Multivariate Generalizations of Cumulative Sum Quality Control Schemes", *Technometrics* **30**, pp. 291-303.

[5] Crowder, S. W., and Hamilton, M. (1992). "Average Run Lengths of EWMA Controls for Monitoring a Process Standard Deviation", *J. Qual. Technol.* **24**, pp. 44-50.

[6] Gumbel, E. J. (1960). "Bivariate Exponential Distributions", *Journal of the American Statistical Association* **55 (2)**, pp. 698-707.

[7] Frank, M. J. (1979). "On the Simultaneous Associativity of F(x, y) and x+y - F(x, y)", *Aequationes Math* **19**, pp. 194-226.

[8] Han, D., Tsung, F., and Li, Y. (2007). "A CUSUM Chart with Local Signal Amplification for Detecting a Range of Unknown Shifts", *International Journal of Reliability* **14(2)**, pp. 81-97.

[9] Hawkins, D. M., Choi, S., and Lee, S. (2007). "A General Multivariate Exponentially Weighted Moving Average Control Chart", *Journal of Quality Technology* **39 (2)**, pp. 118-125.

[10] Hawkins, D. M., and Maboudou Tchao, E. M. (2008). "Multivariate Exponentially Weighted Moving Covariance Matrix", *Technometrics* **50 (2)**, pp. 155-166.

[11] Hotelling, H. (1947). "Multivariate Quality Control Illustrated by the Air Testing of Sample Bombsights," in *Techniques of Statistical Analysis,* eds. C. Eisenhart, M. W. Hastay, and W. A. Wallis, New York: McGraw-Hill, pp. 111-184.

[12] Huh, I. (2010). "Multivariate EWMA Control Chart and Application to a Semiconductor Manufacturing Process," unpublished M.Sc. thesis, McMaster University, Department of Mathematics and Statistics.

[13] Karlin, S., and Taylor, M. H. (1975). *A First Course In Stochastic Processes,* 2nd Edition, Academic Press Inc, New York.

[14] Lowry, C. A., Woodall, W. H., Champ, C. W., and Ridgon, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart", *Technometrics* **34**, pp. 46-53.

[15] Lu, J., and Bhattacharyya, G. K. (1991b). "Inference Procedures for a Bivariate Exponential Distributions", *Statistics & Probability Letters* **12**, pp. 37-50.

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

[17] Lucas, J. M., and Saccucci, M. S. (1990). "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements", *Technometrics* **32**, pp. 1-22.

[18] 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**, pp. 933-949.

[19] Montgomery, DC (2013). *Introduction to Statistical Quality Control,* 7th Edition, Wiley, New York.

[20] Nelsen, R. B. (1999). "An Introduction to Copulas, Lectures Notes in Statistics", **139**, Springer-Verlag, New York.

[21] Pignatiello, J. J., Jr., and Runger, G. C. (1990). "Comparisons of Multivariate CUSUM Charts", *Journal of Quality Technology* **22**, pp. 173-186.

[22] Prabhu, S. S., and Runger, G. C. (1997). "Designing a Multivariate EWMA Control Chart", *Journal of Quality Technology* **29 (1)**, pp. 8-15.

[23] Reynolds, M., and Stoumbos, Z. (2008). "Combination of Multivariate Shewhart and MEWMA Control Charts for Monitoring the Mean Vector and Covariance Matrix", *Journal of Quality Technology* **40 (4)**, pp. 381-393.

[24] Rhoads, T. R., Montgomery, D. C., and Mastrangelo, C. M. (1996). "Fast Initial Response Scheme for the EWMA Control Chart", *Quality Engineering* **9 (2)**, pp. 317-327.

[25] Rigdon, S. E. (1995a). "A Double-Intergal Equation for the Average Run Length of a Multivariate Exponentially Weighted Moving Average Control Chart", *Statistics and Probability Letters* **24**, pp. 365-373.

[26] Rigdon, S. E. (1995b). "An Integral Equation for the In-Control Average Run Length of a Multivariate Exponentially Weighted Moving Average Control Chart", *Journal of Statistical Computation and Simulation* **52**, pp. 351-365.

[27] Roberts, S. W. (1959). "Control chart tests based on geometric moving averages", *Technometrics* **1**, pp. 239-250.

[28] Runger, G. C., and Prabhu, S. S. (1996). "A Markov Chain Model for the Multivariate Exponentially Weighted Moving Averages Control Chart", *Journal of the American Statistical Association* **91**, pp. 1701-1706.

[29] Ryan, T.P. (2000). *Statistical Methods for Quality Improvement,* 2nd Edition, Wiley, New York.

[30] Ryu, J., Wan, H., and S, Kim. (2010). "Optimal Design of a CUSUM Chart for a Mean Shift of Unknown Size", *Journal of Quality Technology* **42**, pp. 311-325.

[31] Schepsmeier, U., Stoeber, J., and Brechmann, C. E. (2013). *VineCopula: Statistical inference of vine copulas.* R package version 1.2. http://CRAN.R-project.org/package=VineCopula

[32] Shamma, S. E., Amin, R. W., and Shamma, A. K. (1991). "A Double Exponentially Weighted Moving Average Control Procedure with Variable Sampling Intervals", *Communications in Statistics - Simulation and Computation* **20**, pp. 511-528.

[33] Shamma, S. E., and Shamma, A. K. (1992). "Development and Evaluation of Control Charts Using Double Exponentially Weighted Moving Averages", *The International Journal of Quality and Reliability Management* **9**, pp. 18-26.

[34] Shu, L., Jiang, W., and Wu, S (2007). "A One-Sided EWMA Control Chart for Monitoring Process Means", *Communications in Statistics - Simulation and Computation* **36**, pp. 901-920.

[35] Sklar, A. (1959). "Fonctions de repartition à n dimensions et leurs marges", *Publications de l'Institut de Statistique de l'Université de Paris.* **8**, pp. 229-231. pp. 901-920.

[36] Steiner, S. H. (1999). "EWMA Control Charts with Time-Varying Control Limits and Fast Initial Response", *Journal of Quality Technology* **31 (1)**, pp. 75-86.

[37] Stoumbos, Z. G., and Sullivan, J, H. (2002). "Robustness to Non-Normality of the Multivariate EWMA Control Chart", *Journal of Quality Technology* **41 (3)**, pp. 217-227.

[38] Woodall, W. H., and Mahmoud, M. A. (2005). "The Inertial Properties of Quality Control Charts", *Technometrics* **47 (4)**, pp. 425-436.

[39] Woodall, W. H., and Maragah, H. D. (1990). Discussion of "Exponentially Weighted Moving Average Control Schemes; Properties and Enhancements ", by J.M Lucas and M.S. Saccucci, *Technometrics* **32**, pp. 17-18.

[40] Woodall, W. H., and Ncube, M. M. (1985). "Multivariate CUSUM Quality Control Procedures", *Technometrics* **27**, pp. 285-292.

[41] Xie, Y., Xie, M., and Goh, T. N. (2011). "Two MEWMA Charts for Gumbel's Bivariate Exponential Distribution", *Journal of Quality Technology* **43 (1)**, pp. 50-65.

[42] Yashchin, E. (1987). "Some Aspects of the Theory of Statistical Control Schemes", *IBM Journal of Research and Development* **31**, pp. 199-205.

[43] Yashchin, E. (1993). "Statistical Control Schemes: Methods, Applications and Generalizations", *International Statistical Review* **61**, pp. 41-66.

[44] Yumin, L. (1996). "An Improvement for MEWMA in Multivariate Quality Control", *Computers in Industrial Engineering* **31**, pp. 779-781.

[45] Zhang, L., and Chen, G. (2005). "An Extended EWMA Mean Chart", *Quality Technology and Quantitative Management* **2**, pp. 39-52.

[46] Zhao, Y., Tsung, F., and Wang, Z. (2005). "Dual CUSUM Control Schemes for Detecting a Range of Mean Shifts", *IIE Transcations* **37**, pp. 1047-1057.

[47] Zou, C., and Tsung, F. (2008). "Directional MEWMA Schemes for Multistage Process Monitoring and Diagnosis", *Journal of Quality Technology* **40 (4)**, pp. 407-427.