A STUDY ON THE PERFORMANCES OF UNIVARIATE AND MULTIVARIATE SYNTHETIC CHARTS

SARANYA THANGARAJAH

MASTER OF MATHEMATICS

LEE KONG CHIAN FACULTY OF ENGINEERING AND SCIENCE

UNIVERSITI TUNKU ABDUL RAHMAN

APRIL 2017

A STUDY ON THE PERFORMANCES OF UNIVARIATE AND MULTIVARIATE SYNTHETIC CHARTS

By

SARANYA THANGARAJAH

A project report submitted to the Department of Mathematical and

Actuarial Sciences,

Lee Kong Chian Faculty of Engineering and Science,

Universiti Tunku Abdul Rahman,

in partial fulfillment of the requirements for the degree of Master of Mathematics

in April 2017

DECLARATION

I, Saranya Thangarajah hereby declare that this project report is based on my original work except for the citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

____

(SARANYA THANGARAJAH)

____

Date

APPROVAL SHEET

The project report entitled "**A STUDY ON THE PERFORMANCES OF UNIVARIATE AND MULTIVARIATE SYNTHETIC CHARTS"** was prepared by SARANYA THANGARAJAH and submitted as a partial fulfillment of the requirements for the degree of Master of Mathematics at Universiti Tunku Abdul Rahman.

Approved By:

…………………………. Date: ……………………

DR. WONG VOON HEE

Assistant Professor / Supervisor

Head of Programme (Master of Mathematics)

Department of Mathematical and Actuarial Sciences,

Lee Kong Chian Faculty of Engineering and Science,

University Tunku Abdul Rahman

ABSTRACT

The Average Run Length (ARL) has been traditionally used as the sole measure of a chart's performance because of the relative difficulty in computing the run length distribution and that the in-control run length distribution is approximately geometric, hence, it can be approximately characterized by the ARL. However, many have criticized the sole use of ARL as a measure of a chart's performance. Therefore, in this project, we have compared the usage of the Average Run Length (ARL), Median Run Length (MRL) and Standard Deviation Run Length (SDRL) in both cases, the univariate and the multivariate.

In this project, we aim to achieve two main objectives. Firstly, to develop programs to compute the average run length (ARL), median run length (MRL) and standard deviation of run length (SDRL) of univariate and multivariate synthetic charts and to compare the performances of univariate and multivariate synthetic charts with the existing control charts.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr Wong Voon Hee for offering me an opportunity to do this project and providing me with valuable guidance and support at every stage of this project. I thank him from the bottom of my heart. Besides that, I would also like to thank my family and friends who have given me their ceaseless support, encouragement and motivation throughout this study and completion of this project.

TABLE OF CONTENTS

		PAGE
DECLARATION		**ii**
APPROVAL SHEET		**iii**
ABSTRACT		**iv**
ACKNOWLEDGEMENT		**v**
TABLE OF CONTENTS		**vi**
LIST OF TABLES		**ix**
LIST OF FIGURES		**x**
LIST OF SYMBOLS/ABBREVIATIONS		**xi**
CHAPTER		
1	**PRELIMINARY STUDY ON SOME UNIVARIATE CONTROL CHARTS**	**1**
	1.1 Control Charts	1
	1.2 $\bar{X}$ Chart	4
	1.3 Conforming Run Length (CRL) Chart	6
	1.4 Performance Measure of Control Chart	6
	1.5 Synthetic Control Chart	7
	1.6 Exponentially Weighted Moving Average (EWMA) Control Chart	10

2	**PERFORMANCES OF SYNTHETIC $\bar{X}$ AND EWMA CHARTS**	**13**
	2.1 ARL, SDRL and MRL Performances of the Synthetic Chart	13
	2.2 ARL, SDRL and MRL Performances of the EWMA Chart	14
	2.3 Comparison of the ARL, MRL and SDRL Performances of the Synthetic $\bar{X}$ and EWMA Charts	16
	2.3.1 Comparison of the ARL Performances of the Synthetic $\bar{X}$ and EWMA Charts	16
	2.3.2 Comparison of the SDRL Performances of the Synthetic $\bar{X}$ and EWMA Charts	17
	2.3.3 Comparison of the MRL Performances of the Synthetic $\bar{X}$ and EWMA Charts	18
	2.3.4 Comparison of the ARL, MRL and SDRL Performances of the Synthetic $\bar{X}$ and EWMA Charts (Conclusion)	19
3	**MULTIVARIATE ANALYSIS**	**23**
	3.1 Introduction	23
	3.2 The Multivariate Normal Distribution	24
	3.3 Multivariate Hotelling T^2 Control Chart	25
	3.4 Multivariate Synthetic T^2 Control Chart	27
	3.5 Multivariate Exponentially Weighted Moving Average (MEWMA) Control Chart	29

4	**PERFORMANCES OF MULTIVARIATE SYNTHETIC AND MEWMA CHARTS**	**30**
	4.1 ARL, SDRL and MRL Performances of the Multivariate Synthetic Chart	30
	4.2 ARL, SDRL and MRL Performances of the MEWMA Chart	31
	4.3 Comparison of the ARL, MRL and SDRL Performances of the Multivariate Synthetic and MEWMA Charts	33
	4.3.1 Comparison of the ARL Performances of the Multivariate Synthetic and MEWMA Charts	33
	4.3.2 Comparison of the SDRL Performances of the Multivariate Synthetic and MEWMA Charts	34
	4.3.3 Comparison of the MRL Performances of the Multivariate Synthetic and MEWMA Charts	35
	4.3.4 Comparison of the ARL, MRL and SDRL Performances of the Multivariate Synthetic and MEWMA Charts (Conclusion)	36
5	**CONCLUSION**	**40**
REFERENCES		**42**
APPENDICES		**44**
	A UNIVARIATE ANALYSIS	44
	B MULTIVARIATE ANALYSIS	48

LIST OF TABLES

TABLE	TITLE	PAGE
2.1	ARLs, SDRLs and MRLs of the Synthetic Chart for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$	13
2.2	ARLs, SDRLs and MRLs of the EWMA Chart for n = 1, 4 and 10, based on $\text{ARL}_0 = 370$	14
2.3	ARLs of the Synthetic $\bar{X}$ and EWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$	16
2.4	SDRLs of the Synthetic $\bar{X}$ and EWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$	17
2.5	MRLs of the Synthetic $\bar{X}$ and EWMA charts for n = 1, 4 and 10 based on $\text{MRL}_0 = 370$	18
4.1	ARLs, SDRLs and MRLs of the Multivariate Synthetic Chart for $n = 1, 4$ and 10 based on $\text{ARL}_0 = 370$	30
4.2	ARLs, SDRLs and MRLs of the MEWMA Chart for $n = 1, 4$ and 10 based on $\text{ARL}_0 = 370$	31
4.3	ARLs of the Multivariate Synthetic and MEWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$	33
4.4	SDRLs of the Multivariate Synthetic and MEWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$	34
4.5	MRLs of the Multivariate Synthetic and MEWMA charts for n = 1, 4 and 10 based on $\text{MRL}_0 = 370$	35

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	A typical control chart	4
1.2	$\bar{X}$ chart	5
1.3	CRL chart	6
1.4	EWMA chart	12
2.1	ARL comparison for the Synthetic Chart and the EWMA Chart	20
2.2	SDRL comparison for the Synthetic Chart and the EWMA Chart	21
2.3	MRL comparison for the Synthetic Chart and the EWMA Chart	22
4.1	ARL comparison for the Multivariate Synthetic and MEWMA Charts	37
4.2	SDRL comparison for the Multivariate Synthetic and MEWMA Charts	38
4.3	MRL comparison for the Multivariate Synthetic and MEWMA Charts	39

LIST OF SYMBOLS/ABBREVIATIONS

SPC	Statistical Process Control
CL	Center Line
UCL	Upper Control Limit
LCL	Lower Control Limit
CRL	Conforming Run Length
ARL	Average Run Length
MRL	Median Run Length
SDRL	Standard Deviation Run Length
EWMA	Exponentially Weighted Moving Average

CHAPTER 1

PRELIMINARY STUDY ON SOME UNIVARIATE CONTROL CHARTS

1.1 Control Charts

Control charts fall under one of the seven major tools of the Statistical Process Control (SPC) which is a powerful collection of problem-solving tools. SPC is greatly useful in attaining process stability and improving capability by the reduction of variability. It is an industry standard methodology for measuring and controlling quality during the manufacturing process. Apart from that, it helps in determining whether a process is in control. The seven major tools which are known as ‘the magnificent seven are briefly described as follows:-

1. Histogram or stem-and-leaf plot

   A histogram is a diagram consisting of vertically placed rectangles whose area is relative to the frequency of a variable and whose width is equal to the class interval. Meanwhile, a stem and leaf plot is an extraordinary table where each data value is separated into a ‘stem’ (the first digit or digits) and a ‘leaf’ (usually the last digit).

2. Check sheet

   A check sheet or sometimes referred to as a tally sheet is a form that is useful in collecting data in real time at the location where the data is produced.

3. Pareto chart

   A Pareto chart consists of both bars and a line graph. In this chart, individual values are denoted in a downward order by bars and the cumulative total is represented by the line.

4. Cause-and-effect diagram

   A cause-and-effect diagram which is also called an Ishikawa diagram or a fishbone diagram is a visualization tool for classifying the prospective causes of a problem in order to determine its root causes.

5. Defect concentration diagram

   A defect concentration diagram also known as a problem concentration diagram is a graphical tool that is beneficial in studying the causes of the product or part defects.

6. Scatter diagram

   A scatter diagram is a graph in which the values of two variables are plotted along two axes and the pattern of the resulting plot of points reveals any correlation that is present between the two variables.

7. Control chart

   A control chart is a graph that is used to study and observe how a process changes over time.

Among these seven tools, the control chart which was invented by Dr Walter A. Shewhart in 1924, turns out to be the most technically sophisticated tool. In simple words, a control chart is very much similar to a patient's temperature chart which tells the doctor the progress of a patient's condition. Likewise, a control chart gives a picture of the continuing story of the state of quality of the manufactured items. These control charts are very good method for problem solving and to improve quality based on the analysis and results obtained as they are used for the purpose of detecting assignable causes that affect process stability. Apart from that, they are also excellent decision makers in leading us in deciding if the process is a good one, a poor one or has no effect based on the pattern of the plotted points no. The control charts are widely used in the industries due to five reasons: proven technique for improving productivity, effective in defect prevention, prevent unnecessary process adjustment, provide diagnostic information and provide information about process capability.

A control chart consists of three important lines namely the center line (CL), upper control limit (UCL) and lower control limit (LCL). The center line can have three different clarifications, based on the available data. It can be the average of the plotted points or it can a reference value and last but not least, it may also represent the population mean, μ , if this value is known. Meanwhile, the UCL and LCL are established to assist in judging the importance of the difference in the quality of the product or service (Besterfield, 2013). These limits are also known as 'action limits' because if any point is plotted beyond these limits, the process is considered to be out of control. Therefore, proper investigation and corrective action must be taken to identify and eradicate the assignable cause(s) which has led to this behaviour. A typical control chart is illustrated in Figure 1.1.

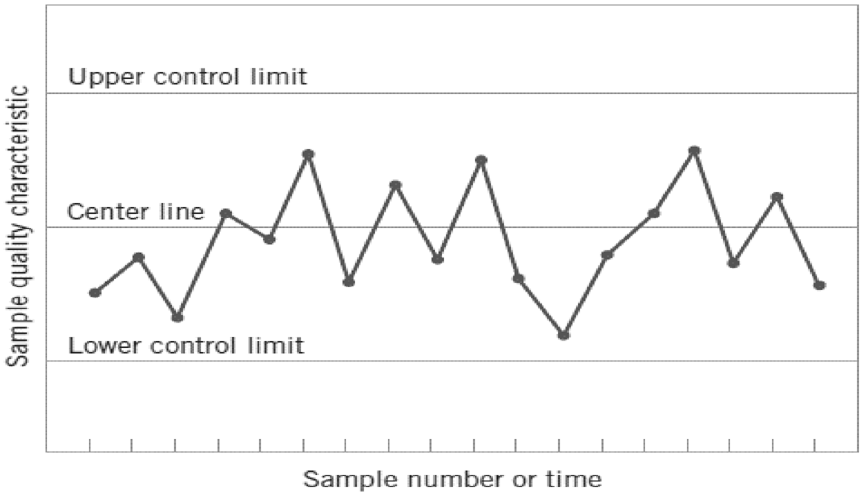

Figure 1.1 A typical control chart

1.2 $\bar{X}$ Chart

The $\bar{X}$ chart is widely used in the industry to monitor the arithmetic means of successive samples of constant size, *n*. The sample means are assumed to be normally distributed, for the purpose of control limit calculation. In an $\bar{X}$ chart, the vertical axis in a control chart denotes the sample average, $\bar{x}$. The process mean is said to be in control if the value of $\bar{x}$ plots between the UCL and LCL lines. Otherwise, the process mean is considered out of control. When a process is in statistical control, the mean value for each subgroup is steady over time and the variation within a subgroup is also steady. This chart is quite effective in detecting large scale shifts in the mean but not so effective in detecting small or moderate shifts in the mean. The general model of a control chart is as follows (on the next page):-

Let w be a sample statistic that measures some quality characteristic of interest and the mean of w is μ_w and the standard deviation of w is σ_w. The center line, upper control and lower control limits are calculated suing the following equations:-

$$\text{UCL} = \mu_w + k\sigma_w \tag{1.1a}$$

$$\text{CL} = \mu_w \tag{1.1b}$$

$$\text{LCL} = \mu_w - k\sigma_w \tag{1.1c}$$

where the factor k controls the width of the limits. Figure 1.2 shows a typical $\bar{X}$ chart used to monitor the weekly average bowling scores.

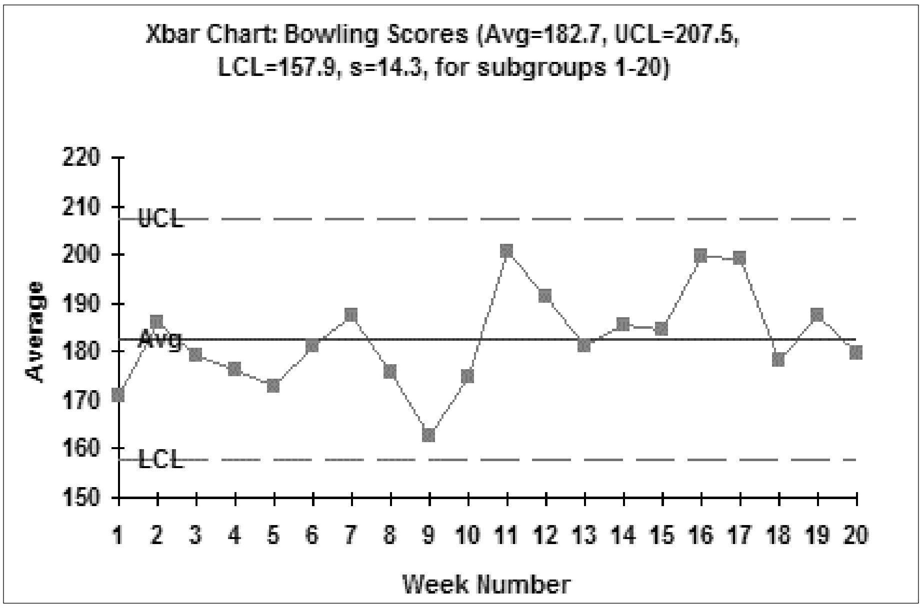

Figure 1.2: $\bar{X}$ chart

1.3 Conforming Run Length (CRL) Chart

The CRL value represents the number of scrutinised samples between two consecutive non-conforming samples including the nonconforming sample in the end where the line is drawn. The hollow circle (white, empty circle) represents conforming samples and the black circle represents non-conforming samples. An example of a CRL chart is illustrated below where $\mathrm{CRL}_2 = 5$ consisting of four conforming samples (hollow circles) between the two non-conforming samples (dark circles) while $\mathrm{CRL}_3 = 3$ consists of two conforming samples between the non-conforming samples.

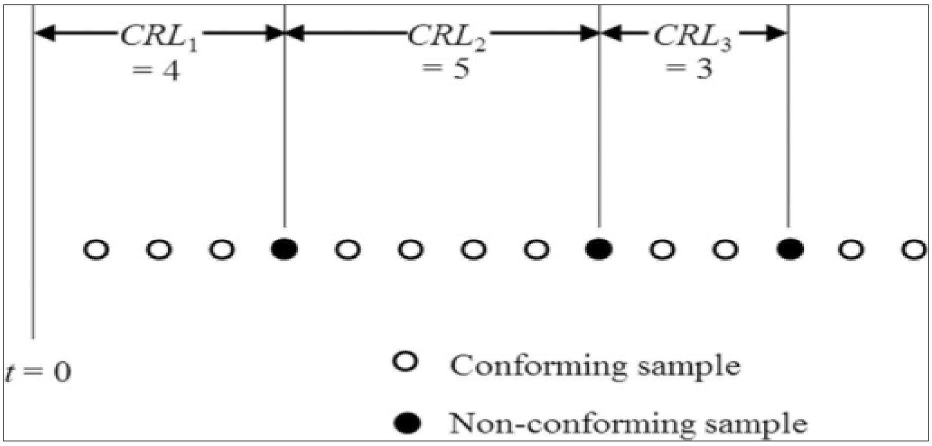

Figure 1.3: CRL chart

1.4 Performance Measure of Control Chart

The performance of control charts in this project is measured in terms of Average Run Length (ARL), Median Run Length (MRL) and Standard Deviation Run Length (SDRL). The ARL is the average number of sample points that are plotted on a chart before the first out-of-control signal is detected while the MRL denotes the median number of sample points that are plotted on a chart before the first out-of-control signal is detected. Last not least, SDRL measures the spread of the run length distribution.

1.5 Synthetic Control Chart

A combination of Shewhart $\bar{X}$ and CRL charts produces a synthetic control chart. Therefore, this synthetic chart contains $\bar{X}/S$ sub-chart and CRL/S sub-chart. Following are the steps involved in constructing and executing the synthetic chart (Wu and Spedding, 2000a):-

Step 1 : The upper and lower control limits of the $\bar{X}/S$ sub-chart are computed according to the formulas below:-

$$\mathrm{UCL}_{\bar{X}/S} = \mu + k\sigma_{\bar{X}} \tag{1.2a}$$

$$\mathrm{LCL}_{\bar{X}/S} = \mu - k\sigma_{\bar{X}} \tag{1.2b}$$

where

- μ - the in-control process mean
- $\sigma_{\bar{X}} = \frac{\sigma_0}{\sqrt{n}}$, standard deviation of the sample mean $\bar{X}$ and n is the sample size
- Factor k - controls the width of the limits, $k > 0$

Step 2 : The lower control limit, L, of the CRL/S sub-chart is to be computed.

Step 3 : A random sample of n observations, X_i, for $i = 1, 2, ..., n$ is to be taken at each inspection point and the sample mean, $\bar{X}$ is calculated.

Step 4 : If $\mathrm{LCL}_{\bar{X}/S} < \bar{X} < \mathrm{UCL}_{\bar{X}/S}$, the sample will be denoted as a conforming sample and the control flow returns to Step 3. Otherwise, the sample is non-conforming and the control flow will continue to Step 5.

Step 5 : The CRL/S sub-chart is used in this step. The number of samples between the current and the last non-conforming is to be tallied. This number is then taken as the CRL value of the CRL/S sub-chart in the synthetic chart.

Step 6 : If this CRL value is larger than the lower control limit, L, of the CRL/S sub-chart, the process is in-control and the control flow goes back to Step 3. Otherwise, the synthetic chart will signal an out-of-control condition.

Step 7 : Corrective actions are taken and implemented to identify and remove the assignable cause/causes. The control flow then returns to Step 3.

The ARL of the synthetic control chart is calculated as follows:-

$$\text{ARL} = \frac{1}{P} \times \frac{1}{1-(1-P)^L} \tag{1.3}$$

where

$$P = 1 - [(k - \delta\sqrt{n}) - (-k - \delta\sqrt{n})] \tag{1.4}$$

$\delta = \frac{|\mu_1 - \mu_0|}{\sigma}$, where δ is the size of the standardized mean shift; μ_1 is the out-of-control mean and μ_0 is the in-control mean. The process is said to be in-control when $\delta = 0$.

The optimal design procedure for the synthetic chart, based on MRL is similar to the procedure for the optimal design of the synthetic chart based on ARL as provided in Wu and Spedding (2000). The steps are as stated on the following page:-

Step 1 : Specify n, δ_{opt} and MRL(0).

Step 2 : Initialize $L = 1$.

Step 3 : By fixing $\gamma = 0.5$ and $\delta = 0$, k is determined using the two equations below:-

$$\Pr(M \le m_\gamma - 1) \le \gamma \tag{1.5a}$$

$$\Pr(M \le m_\gamma) > \gamma \tag{1.5b}$$

$$\Pr(M \le m) = \mathbf{s}^T(\mathbf{I} - \mathbf{R}^m)\mathbf{1} \tag{1.5c}$$

where $\mathbf{I}$ is the $(L+1)\times(L+1)$ identity matrix, $\mathbf{1}$ is a vector with each of its $(L+1)$ elements equal to unity and $\mathbf{s}$ is the initial probability column vector having $(L+1)$ elements, with a single element equal to one and zero elsewhere. The value of $m_{0.5} = \text{MRL}(0)$ as specified in Step 1.

Step 4 : Calculate $\text{MRL}(\delta_{\text{opt}})$ from the current k and L values using Equations (1.5a) and (1.5b) where $\text{MRL}(\delta_{\text{opt}}) = m_{0.5}$ when $\delta = \delta_{\text{opt}}$, by setting $\gamma = 0.5$ and $\delta = \delta_{\text{opt}}$.

Step 5 : If $L = 1$, increase L by one and the flow returns to Step 3. If $L \ge 2$ check whether the $\text{MRL}(\delta_{\text{opt}})$ of the current L has been reduced compared to that of $L-1$. If it turns out to be a reduction, L is to be increased by one and the flow returns to Step 3. Otherwise, the process will move on to Step 6.

Step 6 : Record the values of L and k that give the smallest $\text{MRL}(\delta_{\text{opt}})$ value as the optimal values for the synthetic chart.

Step 7 : The optimal value k in Step 6 will be used to compute $\mathrm{UCL}_{\bar{X}/S}$ and $\mathrm{LCL}_{\bar{X}/S}$. The optimal value L will be used as the lower limit of the CRL/S sub-chart.

1.6 Exponentially Weighted Moving Average (EWMA) Control Chart

The EWMA control chart was introduced by Roberts (1959) and further explored by Crowder (1987a, 1989) and Lucas and Saccucci (1990). It is a time-weighted control chart that plots exponentially weighted moving averages. The EWMA chart is used in statistical process control to monitor variables that make use of the total history of a given output.

The EWMA statistic is calculated as follows:-

$$Z_i = \lambda \bar{X}_t + (1-\lambda) Z_{i-1} \tag{1.6}$$

where $0 < \lambda \le 1$ is a constant and the starting value (required with the first sample at $i = 1$) is the process target whereby $z_0 = \mu_0$. Z_i is a weighted average of all previous sample means. The center line and control limits for the EWMA control chart are defined as follows:-

$$\mathrm{UCL} = \mu_0 + L\sigma\sqrt{\frac{\lambda}{n(2-\lambda)}[1-(1-\lambda)^{2i}]} \tag{1.7a}$$

$$\mathrm{CL} = \mu_0 \tag{1.7b}$$

$$\mathrm{LCL} = \mu_0 - L\sigma\sqrt{\frac{\lambda}{n(2-\lambda)}[1-(1-\lambda)^{2i}]} \tag{1.7c}$$

where the factor L controls the width of the control limits.

The control flow for the EWMA is more simplified and shorter compared to the synthetic control chart. In a EWMA chart, when Z_i plots beyond the limits, above the UCL or below the LCL, an out-of-control is signalled.

The ARL of the EWMA control chart is calculated as follows:-

$$\text{ARL} = \boldsymbol{u}'(\mathbf{I} - \boldsymbol{R})^{-1}\mathbf{1} \qquad (1.8)$$

where $\boldsymbol{u} = (0,...,0,1,0,...,0)'$, is the initial probability vector.

The SDRL of the EWMA control chart is calculated as follows:-

$$\text{SDRL} = \sqrt{2\boldsymbol{u}'(\mathbf{I} - \boldsymbol{R})^{-2}\boldsymbol{R}\cdot\mathbf{1} - (\text{ARL})^2 + \text{ARL}} \qquad (1.9)$$

where

$\mathbf{1}$ denotes a vector of ones

$\mathbf{I}$ represents the identity matrix

$\boldsymbol{R}$ is the transition probability matrix for the transient states

Figure 1.2 on the following page demonstrates a typical EWMA chart which is used to monitor the diameter of all rotors produced in a specific week.

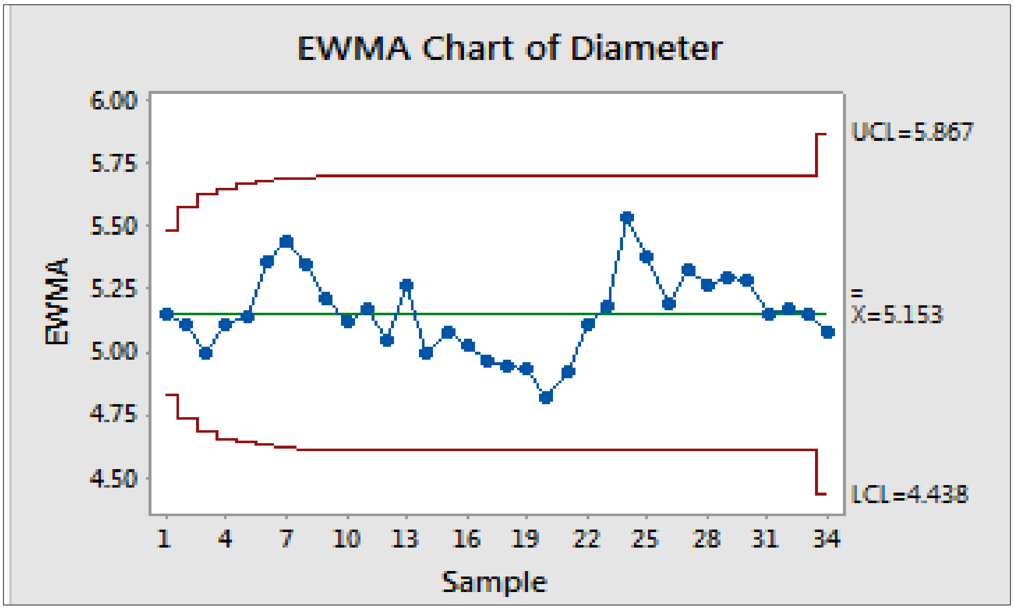

Figure 1.2: EWMA chart

CHAPTER 2

PERFORMANCES OF SYNTHETIC $\bar{X}$ AND EWMA CHARTS

2.1 ARL, SDRL and MRL Performances of the Synthetic Chart

n	**1**	**4**	**10**	**1**	**4**	**10**	**1**	**4**	**10**
L	**19**	**5**	**2**	**19**	**5**	**2**	**9**	**2**	**1**
k	**2.4945**	**2.2601**	**2.0845**	**2.4945**	**2.2601**	**2.0845**	**2.4476**	**2.1649**	**2.0252**
δ		**ARL**			**SDRL**			**MRL**	
0	369.95	369.84	369.84	433.63	406.55	394.00	370	370	370
0.25	253.34	122.84	52.77	303.21	141.73	60.28	250	126	54
0.5	109.22	22.61	6.25	136.90	28.22	7.47	101	22	6
0.75	44.23	6.23	1.92	57.35	7.55	1.76	36	5	1
1	20.04	2.73	1.19	25.58	2.66	0.56	9	2	1
1.5	6.49	1.30	1.00	6.70	0.63	0.06	4	1	1
2	3.22	1.04	1.00	2.70	0.21	0.00	2	1	1

Table 2.1: ARLs, SDRLs and MRLs of the Synthetic Chart for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$

Table 2.1 shows the ARL, SDRL and MRL profiles for the synthetic chart based on the magnitude of shift, $0 \le \delta \le 2$. The limit of CRL sub-chart, *L* and the constant *k*, which controls the width of limits for $\bar{X}$ sub-chart are obtained by fixing the in-control ARL, $\text{ARL}_0 = 370 / \text{MRL}_0 = 370$ and the sample sizes at $n = 1, 4 \text{ and } 10$. Two different Mathematica programs were designed to compute the values of ARL, MRL and SDRL. ARL and SDRL were computed using the program in Appendix A.1 while MRL was computed using the program in Appendix A.2.

Overall, the results show that as δ is increasing, the ARLs, SDRLs and MRLs are decreasing. When the value of δ increases from 0 to 0.75, the drop in the values of ARL, MRL and SDRL are faster compared to the increase of δ values from 1 to 2.

2.2 ARL, SDRL and MRL Performances of the EWMA Chart

n	**1**	**4**	**10**	**1**	**4**	**10**	**1**	**4**	**10**
L	**0.4**	**0.391**	**0.712**	**0.4**	**0.391**	**0.712**	**0.634**	**0.513**	**0.81**
λ	**0.05**	**0.145**	**0.722**	**0.05**	**0.145**	**0.722**	**0.096**	**0.211**	**0.81**
δ	**ARL**			**SDRL**			**MRL**		
0	370.69	371.35	370.14	357.48	365.68	369.29	363	363	366
0.25	73.64	31.50	42.41	58.45	24.88	41.14	78	32	46
0.5	26.61	9.61	7.19	15.36	5.06	5.97	25	9	7
0.75	15.40	5.45	2.76	6.95	2.14	1.74	14	5	3
1	10.78	3.84	1.64	4.07	1.24	0.80	9	3	1
1.5	6.78	2.50	1.05	2.00	0.63	0.23	6	2	1
2	5.00	1.98	1.00	1.24	0.38	0.03	4	2	1

Table 2.2: ARLs, SDRLs and MRLs of the EWMA Chart for n = 1, 4 and 10, based on $\text{ARL}_0 = 370$

The optimal parameters L and λ are obtained by minimizing the out-of-control ARL for a desired shift of interest, $\delta_{\text{opt}} = 1$, so that an in-control ARL

($\text{ARL}_0 = 370$) is achieved, based on the sample sizes n = 1, 4 and 10. The value of L, factor that controls the width of the control limits and the constant λ, are obtained by fixing the $\text{ARL}_0 = 370/\text{MRL}_0 = 370$ and the sample sizes at n = 1, 4 and 10. The same method was applied to obtain the results for the EWMA chart, whereby two different Mathematica programs were designed to compute the values of ARL, MRL and SDRL. ARL and SDRL were computed using the program in Appendix A.3 while MRL was computed using the program in Appendix A.4.

Based on the results shown in Table 2.2, it can be studied that as the value of δ is increasing, the values of ARLs, SDRLs and MRLs are decreasing. This clearly shows that the larger the process shift, the easier the out-of-control signals are detected and will result in a smaller spread in the run length distribution. It can also be observed from Table 2.2 that as the sample size increases, the ARL values decrease and this implies that the detection of out-of-control signals is much faster. The same trend is also observed for SDRL whereby the variation of the run length is reduced as the sample size increases. The MRL on the other hand, also reduces as the sample size increases and this implies that the detection process of out-of-control increases.

2.3 Comparison of the ARL, MRL and SDRL Performances of the Synthetic $\bar{X}$ and EWMA Charts

2.3.1 Comparison of the ARL Performances of the Synthetic $\bar{X}$ and EWMA Charts

n	1	4	10	1	4	10
δ	Synthetic			EWMA		
0	369.95	369.84	369.84	370.69	371.35	370.14
0.25	253.34	122.84	52.77	73.64	31.50	42.41
0.5	109.22	22.61	6.25	26.61	9.61	7.19
0.75	44.23	6.23	1.92	15.40	5.45	2.76
1	20.04	2.73	1.19	10.78	3.84	1.64
1.5	6.49	1.30	1.00	6.78	2.50	1.05
2	3.22	1.04	1.00	5.00	1.98	1.00

Table 2.3: ARLs of the Synthetic $\bar{X}$ and EWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$

Table 2.3 presents the comparison between the ARL values for the Synthetic $\bar{X}$ chart and the EWMA chart. For both charts, when the comparison is made horizontally, it can be observed that as the sample size, n increases, the ARL values decrease which also indicates that it is much faster to detect an out-of-control. If observation is made vertically, the ARL values decrease as the value of δ increases. The ARL produced by the EWMA chart is smaller than the values produced by the synthetic chart when $0 < \delta \le 0.75$. On the other hand, when $\delta \ge 1$, EWMA chart recorded larger values of ARL. The smaller ARL values recorded by the EWMA chart when $0 < \delta \leq 0.75$ indicates that the EWMA chart is the preferred chart for detecting small mean shifts. On the hand, the larger values of ARL that were recorded by the EWMA chart when $\delta \geq 1$ indicates that it is not a good option for detecting larger mean shifts. The synthetic chart would be the preferred option in detecting moderate and larger mean shifts.

2.3.2 Comparison of the SDRL Performances of the Synthetic $\bar{X}$ and EWMA Charts

n	**1**	**4**	**10**	**1**	**4**	**10**
δ	**Synthetic**			**EWMA**		
0	433.63	406.55	394.00	357.48	365.68	369.29
0.25	303.21	141.73	60.28	58.45	24.88	41.14
0.5	136.90	28.22	7.47	15.36	5.06	5.97
0.75	57.35	7.55	1.76	6.95	2.14	1.74
1	25.58	2.66	0.56	4.07	1.24	0.80
1.5	6.70	0.63	0.06	2.00	0.63	0.23
2	2.70	0.21	0.00	1.24	0.38	0.03

Table 2.4: SDRLs of the Synthetic $\bar{X}$ and EWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$

The table above compares the performances of the Synthetic $\bar{X}$ chart and the EWMA chart based on the SDRL values that were produced. By studying Table 2.4, it can be observed that as the sample size increases the SDRL values decrease. This indicates that the variation of the run length is reduced as the sample size increases. Small value of SDRL indicates that there is less variation in the run length and it will be easy to estimate the number of samples that will be needed to detect an out-of-control. The performance of the chart is also said to be more stable when the SDRL values are small. Therefore, when compared, the EWMA chart surpasses the synthetic chart for small and moderate δ values, where $\delta \le 1.5$. So, in this case, the EWMA chart is the favoured chart in detecting small and moderate mean shifts and the synthetic chart should be used detecting the larger mean shifts.

2.3.3 Comparison of the MRL Performances of the Synthetic $\bar{X}$ and EWMA Charts

n	**1**	**4**	**10**	**1**	**4**	**10**
δ	**Synthetic**			**EWMA**		
0	370	370	370	363	363	366
0.25	250	126	54	78	32	46
0.5	101	22	6	25	9	7
0.75	36	5	1	14	5	3
1	9	2	1	9	3	1
1.5	4	1	1	6	2	1
2	2	1	1	4	2	1

Table 2.5: MRLs of the Synthetic $\bar{X}$ and EWMA charts for n = 1, 4 and 10 based on $\text{MRL}_0 = 370$

Table 2.5 projects the results of the Synthetic $\bar{X}$ and the EWMA chart when the MRL values are taken into consideration. The same kind of results that were obtained for the ARL observation is seen here. When $0 < \delta \leq 0.75$, the MRL values that were produced by the EWMA chart were smaller than the synthetic chart. But the values observed in the synthetic chart were larger than the values in the EWMA chart for $\delta \geq 1$.The only difference is that the values of MRL tend to converge to a certain value as the value of δ increases. So it can be concluded that the synthetic chart is the preferred chart for detecting moderate and larger mean shifts whereas the EWMA chart is to be used for detecting smaller mean shifts.

2.3.4 Comparison of the ARL, MRL and SDRL Performances of the Synthetic $\bar{X}$ and EWMA Charts (Conclusion)

Based on Wu and Spedding (2000), it has been concluded that the synthetic chart outperforms the $\bar{X}$ chart. Therefore, in this dissertation, only the synthetic chart and the EWMA charts are compared to obtain results that were seen in Sections 2.1, 2.2 and 2.3. In a nutshell, the EWMA chart should be used to monitor a process if past experiences indicates that a small mean shift is likely to occur. On the other hand, if historical data demonstrate that a moderate or a large shift usually happens, then the synthetic chart should be the preferred chart that is to be used in process monitoring.

Figures 2.1, 2.2 and 2.3 show the differences between the ARL, SDRL and MRL values for the Synthetic $\bar{X}$ the EWMA charts, respectively.

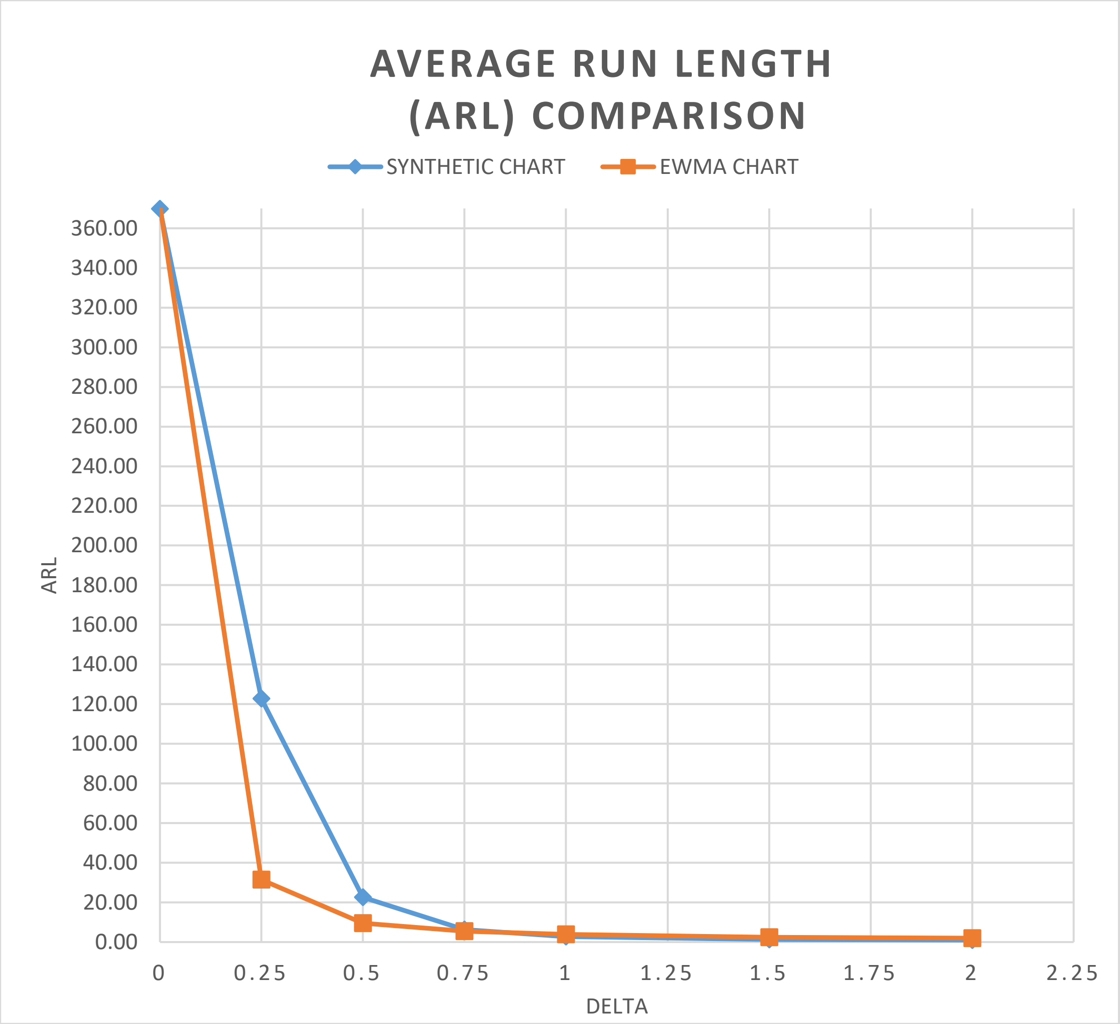

Figure 2.1: ARL comparison for the Synthetic Chart and the EWMA Chart

The line chart above clearly shows that the EWMA chart is the preferred chart for small mean shifts. However, the ARL chart is the preferred chart for moderate and larger shifts.

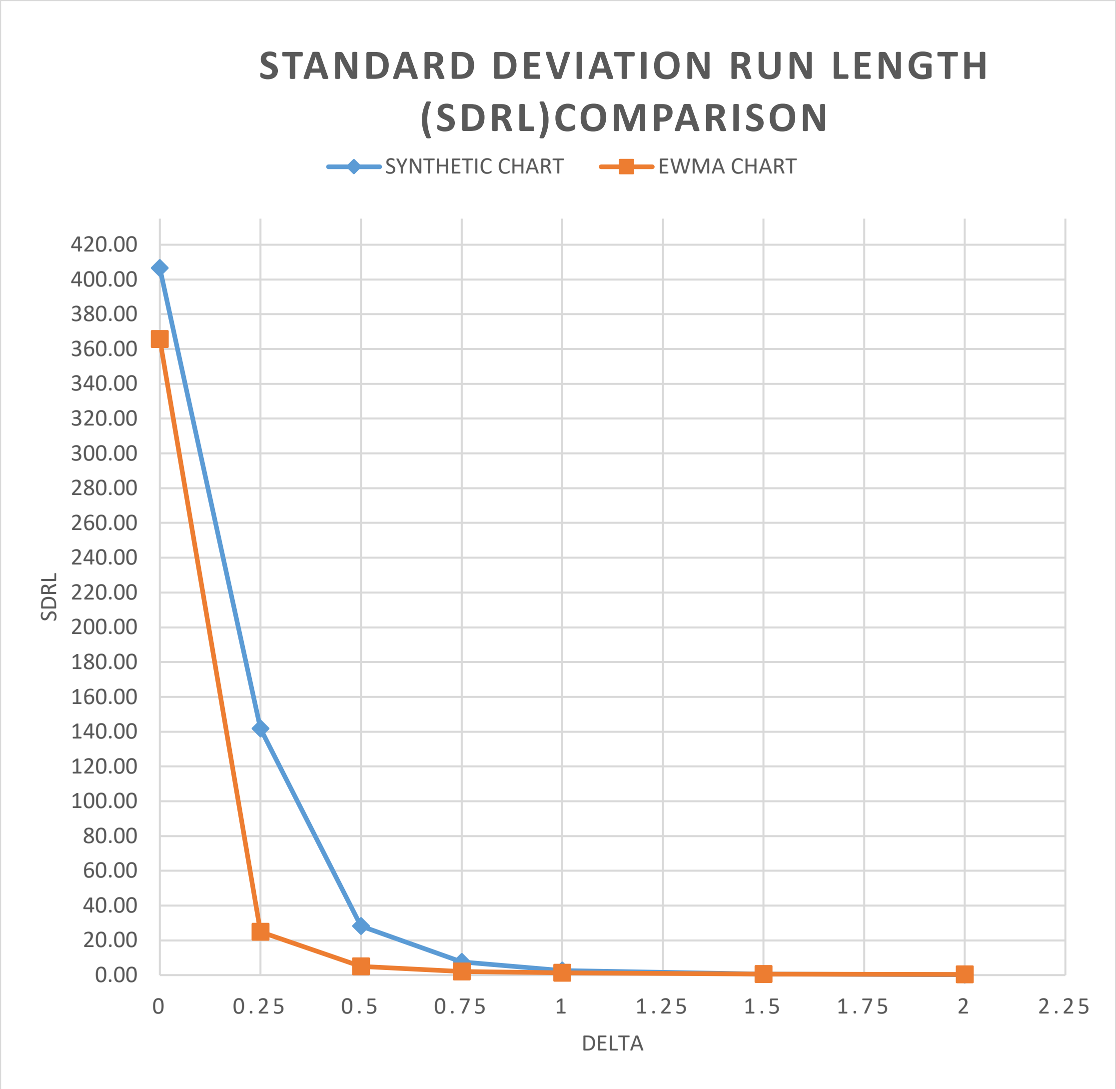

Figure 2.2: SDRL comparison for the Synthetic Chart and the EWMA Chart

The similar result is obtained here whereby the EWMA chart outperforms the synthetic chart for $\delta \leq 1.5$ and this indicates that the EWMA chart is the preferred chart for detecting small and also moderate mean shifts.

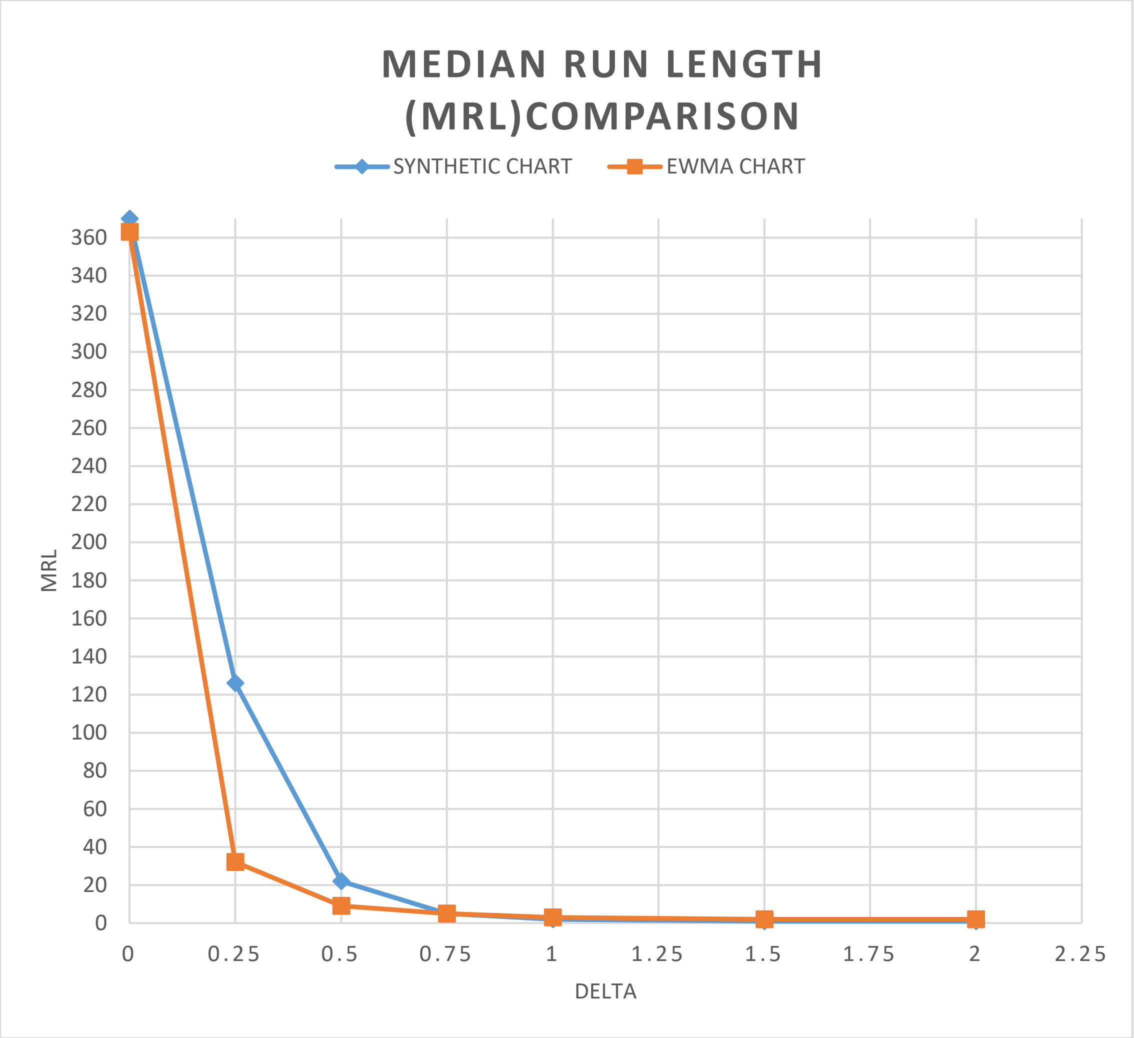

Figure 2.3: MRL comparison for the Synthetic Chart and the EWMA Chart

As for the MRL, it can also be seen that the EWMA is the preferred chart when the shift in mean is small. Therefore, it can be concluded that the EWMA is a better choice for small mean shifts, meanwhile, the synthetic chart should be the favoured option for moderate and large shifts.

CHAPTER 3

MULTIVARIATE ANALYSIS

3.1 Introduction

Multivariate analysis, as the name indicates, is the analysis of data sets with more than one variable. It is a statistical process of instantaneously analysing numerous independent variables with numerous dependent variables. This is more realistic compared to the univariate analysis as it models reality where each situation, product or decision in most of our daily lives involves more than a single variable. Therefore, multivariate analysis can be used to process the information in a meaningful way when available information and data are stored in database tables containing rows and columns.

Multivariate analysis methods are commonly used in a variety of industries, such as the consumer and research market, process optimization and process control, research and development and last but not least quality control and quality assurance across a range of industries such as food and beverage, paint, pharmaceuticals, chemicals, energy and telecommunications. If these data were to be monitored and analysed with univariate SPC procedures, it will often turn out to be ineffective as the total size of these databases is usually measured in millions of individual records. For this reason, the use of multivariate methods has increased tremendously in recent years.

3.2 The Multivariate Normal Distribution

The same approach that was used in the univariate statistical control is applied to the multivariate normal distribution. Suppose there are p variables denoted as $x_1, x_2, ..., x_p$. These variables are arranged in a p-component vector $\mathbf{x}' = [x_1, x_2, ..., x_p]$. Let the vector of the means of the x's be $\mu' = [\mu_1, \mu_2, ..., \mu_p]$ and let the variances and covariances of the random variables in **x** be contained in a $p \times p$ covariance matrix Σ. The main diagonal elements of Σ are the variances of x and the off-diagonal elements are the covariances. The squared standardized (generalized) distance from **x** to μ is (Montgomery, 2009)

$$(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}). \tag{3.1}$$

The multivariate normal density function is obtained by replacing the standardized distance in the univariate normal distribution, $(x-\mu)(\sigma^2)^{-1}(x-\mu)$ with the multivariate generalized distance in Equation (3.1) and changing the constant term $\frac{1}{\sqrt{2\pi\sigma^2}}$ to a more general form that makes the area under the probability density function unity regardless of the value of p. With this, the multivariate normal probability density function is given by (Montgomery, 2009)

$$\mathbf{f(x)} = \frac{\mathbf{1}}{(\mathbf{2}\boldsymbol{\pi})^{\mathbf{p/2}}\left|\Sigma\right|^{\mathbf{1/2}}}\mathbf{e}^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}, \tag{3.2}$$

where $-\infty < x_j < \infty$, $j = 1, 2, ..., p$.

3.3 Multivariate Hotelling T^2 Control Chart

The original work in multivariate quality control was done by Hotelling (1947), who applied his procedures to bombsight data during World War II. Before this study, he wrote a paper on T^2- test procedures for multivariate population (1931). Successive papers dealing with control procedures for several related variables include Hicks (1995), Jackson (1956, 1959), Edward Jackson (1985), Crosier (1988), Hawkins (1991, 1993), Lowry and Montgomery (1995), Pignatiello and Runger (1990), Tracy (1992), Montgomery and Wadsworth (1972). El-Din et al. (2006) made an evaluation of the univariate and the multivariate out-of-control signals to demonstrate the effectiveness of the Hotelling T^2. Moreover, Haridy and Wu (2009) also made a comparison between the univariate and multivariate quality control charts for observing dynamic behaviour processes.

The Hotelling T^2 control chart is known to be one of the most popular multivariate control charts. It is used for monitoring the mean vector of the process. This control chart is a multivariate extension of the Shewhart $\overline{X}$ chart. There are namely 2 versions of the Hotelling T^2 chart, for the sub-grouped data and for the individual observations.

For the sub-grouped data, suppose that two quality characteristics x_1 and x_2 are jointly distributed according to the bivariate normal distribution. Let μ_1 and μ_2 be the mean values of the quality characteristics and let σ_1 and σ_2 be the standard deviations of x_1 and x_2 respectively. The covariance between x_1

and x_2 is denoted by σ_{12}. We assume that σ_1, σ_2 and σ_{12} are known. If $\bar{x}_1$ and $\bar{x}_2$ are the sample averages of the two quality characteristics calculated from a sample of size *n*, then the statistic

$$\boldsymbol{\chi_0^2 = \frac{n}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}[\sigma_2^2(\bar{x}_1 - \mu_1)^2 + \sigma_1^2(\bar{x}_2 - \mu_2)^2 - 2\sigma_{12}(\bar{x}_1 - \mu_1)(\bar{x}_2 - \mu_2)]} \quad (3.3)$$

will have a chi-square distribution with 2 degrees of freedom. The test statistic plotted on the chi-square control chart for each sample is

$$\boldsymbol{\chi_0^2 = n(\bar{x} - \mu)'\Sigma^{-1}(\bar{x} - \mu)} \quad (3.4)$$

where $\mu' = [\mu_1, \mu_2, ..., \mu_p]$ is the vector of in-control means for each quality characteristic and Σ is the covariance matrix. The upper limit of the control chart is

$$\mathbf{UCL} = \boldsymbol{\chi^2_{\alpha,p}} \quad (3.5)$$

When S is used to estimate Σ and vector is taken as the in-control value of the mean vector of the process, the test statistic for Hotelling T^2 is obtained as follows (μ and Σ is replaced with and S respectively) (Montgomery 2009)

$$\mathbf{T^2 = n(\bar{x} - \bar{\bar{x}})'S^{-1}(\bar{x} - \bar{\bar{x}})} \quad (3.6)$$

For the individual observations, this situation usually occurs where they have multiple quality characteristics that must be looked into, for example, in the chemical and process industries. Assume that there are *m* samples, each of size $n = 1$, are available and the number of quality characteristics that are observed in each sample is *p*. Let $\bar{x}$ and S be the sample mean vector and covariance matrix, respectively. The Hotelling T^2 statistic in equation (3.6) becomes

$$\mathbf{T^2 = (x - \bar{x})'S^{-1}(x - \bar{x})} \tag{3.7}$$

3.4 Multivariate Synthetic T^2 Control Chart

In the univariate section, it was stated that Wu and Spedding developed the synthetic $\overline{X}$ chart which was a combination of the Shewhart $\bar{X}$ chart and the conforming run length (CRL) chart. This was done to boost the performance of the Shewhart $\bar{X}$ chart. We have the same that was done in the multivariate part as well, Ghute and Shirke (2008a) incorporated the Hotelling T^2 and the CRL charts to produce the multivariate synthetic T^2 chart. Thus, the synthetic T^2 chart consists of the T^2 sub-chart and the CRL sub-chart. A CRL value is defined as the number of examined T^2 samples between two sequential non-conforming T^2 samples, including the last non-conforming sample. The synthetic chart is constructed according to the following steps (Ghute amd Shirke, 2008a):-

Step 1 : The upper control limit (UCL) of the T^2 sub-chart and the lower control limit, *L* of the CRL sub-chart is determined. Ghute and Shirke (2008a) and Aparisi and de Luna (2009a) provided formulae to compute the optimal limits of UCL and *L*, by fixing ARL_0 and minimizing the out-of-control ARL, so that the synthetic T^2 has the highest sensitivity for the magnitude of shift of interest, δ_{opt}, considered large enough to seriously weaken the quality of the process. δ_{opt} is measured in terms of the Mahalanobis distance, which is as follows:-

$$\boldsymbol{\delta} = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^{\mathbf{T}} \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}\,, \qquad (3.8)$$

where μ_0 and μ_1 are the in-control and out-of-control mean vectors, respectively and Σ_0 is the in-control covariance matrix.

Step 2 : *p* – variate random observations of size *n* are taken at each sampling point and using the formula below, the T^2 sub-chart's statistic is calculated

$$\mathbf{T}^2 = \mathbf{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)\,, \qquad (3.9)$$

where $\bar{X}$ is the sample mean vector.

Step 3 : The sample mean is considered to be non-conforming if T^2 > UCL. The control flow will then proceed to the next step, Step 4. Otherwise, the sample is categorized as conforming and the control flow will return to Step 2.

Step 4 : The number of T^2 samples between the current (included) and the previous (excluded) non-conforming samples are to be counted and this value will be taken as the CRL value in the CRL sub-chart.

Step 5 : The process is confirmed to be in-control if CRL>L. The control flow will then return to Step 2. Otherwise, it will move on to the final step, Step 6.

Step 6 : An out-of-control is motioned. Corrective actions are to be taken to find and remove the assignable cause(s). Once this is done, the control flow can then return to Step 2.

3.5 Multivariate Exponentially Weighted Moving Average (MEWMA) Control Chart

EWMA control charts were developed to provide more sensitivity to small shifts in the univariate case and they can be extended to multivariate quality control problems. Crosier (1988). Lowry et al. (1992) developed a multivariate EWMA (MEWMA) control chart. It is an extension of the univariate EWMA and is defined as follows

$$\mathbf{Z_i = \lambda x_i + (1 - \lambda) Z_{i-1}} \tag{3.10}$$

where $0 \le \lambda \le 1$ and $Z_0 = 0$. The quantity that is to be plotted on the MEWMA control chart is

$$\mathbf{T_i^2 = Z_i' \Sigma_{Z_i}^{-1} Z_i} \tag{3.11}$$

where the covariance matrix is

$$\mathbf{\Sigma_{Z_i} = \frac{\lambda}{2 - \lambda}[1 - (1 - \lambda)^{2i}] \Sigma} \tag{3.12}$$

CHAPTER 4

PERFORMANCES OF MULTIVARIATE SYNTHETIC AND MEWMA CHARTS

4.1 ARL, SDRL and MRL Performances of the Multivariate Synthetic Chart

n	**1**	**4**	**10**	**1**	**4**	**10**	**1**	**4**	**10**
L	**61**	**26**	**10**	**61**	**26**	**10**	**87**	**16**	**4**
UCL	**9.809**	**9.037**	**8.139**	**9.809**	**9.037**	**8.139**	**10.776**	**9.059**	**7.675**
δ	**ARL**			**SDRL**			**MRL**		
0	369.73	369.66	369.72	463.53	441.08	419.09	370	370	370
0.25	285.37	154.48	67.63	362.68	193.53	83.76	266	141	65
0.5	151.93	34.01	8.19	198.08	44.13	9.92	87	16	4
0.75	72.03	10.36	2.60	93.91	11.46	2.13	53	7	2
1	35.93	4.72	1.46	44.09	4.24	0.82	31	3	1
1.5	12.41	1.77	1.02	12.27	1.17	0.15	12	1	1
2	5.66	1.15	1.00	5.14	0.41	0.01	5	1	1

Table 4.1: ARLs, SDRLs and MRLs of the Multivariate Synthetic Chart for $n = 1, 4$ and 10 based on $\text{ARL}_0 = 370$

Table 4.1 shows the ARL, SDRL and MRL profiles for the multivariate synthetic chart based on the magnitude of shift, $0 \le \delta \le 2$. The upper control limit (UCL) of the T^2 sub-chart and the lower control limit, L of the CRL sub-chart are obtained by fixing the $\text{ARL}_0 = 370 / \text{MRL}_0 = 370$ and the sample sized at $n = 1, 4$ and 10. The computation of the ARL, SDRL and MRL values were obtained from the formation of two Mathematica programs. One program that was designed to compute ARL and SDRL and the other to compute the MRL values. These programs can be found in Appendix B.1 and B.2.

When compared to the univariate section, the results in the multivariate section are very much similar to the univariate section. By studying Table 4.1, the same trend can be observed whereby as the shift in mean, δ, is increasing, the ARLs, SDRLs and MRLs are decreasing. This is a vertical observation. If the observation is done horizontally, it can be concluded that as the sample size increases, from 1 to 10, the values of ARL, SDRL and MRL goes on a downward trend.

4.2 ARL, SDRL and MRL Performances of the MEWMA Chart

n	**1**	**4**	**10**	**1**	**4**	**10**	**1**	**4**	**10**
H	**9.928**	**10.232**	**10.568**	**9.928**	**10.232**	**10.568**	**11.058**	**11.182**	**11.285**
λ	**0.09**	**0.11**	**0.14**	**0.09**	**0.11**	**0.14**	**0.11**	**0.12**	**0.13**
δ	**ARL**			**SDRL**			**MRL**		
0	370.14	369.92	369.93	360.40	361.73	363.32	370	370	370
0.25	110.22	36.66	17.04	98.36	27.54	10.68	109	34	16
0.5	34.66	11.48	5.93	24.39	5.46	2.18	33	11	6
0.75	17.77	6.63	3.66	9.66	2.32	1.01	17	6	4
1	11.66	4.71	2.71	5.13	1.35	0.65	11	5	3
1.5	6.91	3.08	1.97	2.27	0.70	0.29	7	3	2
2	4.96	2.32	1.50	1.35	0.49	0.50	5	2	2

Table 4.2: ARLs, SDRLs and MRLs of the MEWMA Chart for $n = 1, 4$ and 10 based on $\text{ARL}_0 = 370$

Table 4.2 projects the values of ARL, SDRL and MRL for the MEWMA Chart for $0 \le \delta \le 2$. The values of H and λ were generated by fixing

$\mathrm{ARL}_0 = 370 /\ \mathrm{MRL}_0 = 370$ for the sample sizes of $n = 1, 4$ and 10. The Mathematica programs that are attached in Appendix B.3 and B.4 were the ones used to generate the ARL, SDRL and MRL values.

The trend that can be observed here is also similar to the multivariate synthetic chart and also the univariate section. Observing the table in a vertical manner, as δ increases from 0 to 2, the values of ARL, SDRL and MRL decreases. On the other hand, through a horizontal observation, as the sample size increases from 1 to 1o, the values of ARL, SDRL and MRL also decreases. The following section studies the comparison of the ARL for both the multivariate synthetic and the MEWMA charts in a much more in depth manner.

4.3 Comparison of the ARL, MRL and SDRL Performances of the Multivariate Synthetic and MEWMA Charts

4.3.1 Comparison of the ARL Performances of the Multivariate Synthetic and MEWMA Charts

n	**1**	**4**	**10**	**1**	**4**	**10**
δ	Multivariate Synthetic			MEWMA		
0	369.73	369.66	369.72	370.14	369.92	369.93
0.25	285.37	154.48	67.63	110.22	36.66	17.04
0.5	151.93	34.01	8.19	34.66	11.48	5.93
0.75	72.03	10.36	2.60	17.77	6.63	3.66
1	35.93	4.72	1.46	11.66	4.71	2.71
1.5	12.41	1.77	1.02	6.91	3.08	1.97
2	5.66	1.15	1.00	4.96	2.32	1.50

Table 4.3: ARLs of the Multivariate Synthetic and MEWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$

The table above shows the comparison between the multivariate synthetic and the MEWMA charts based on the ARL performances. The ARL values produced by the MEWMA chart are smaller compared to the values in the multivariate synthetic chart when $0 < \delta \leq 1$. Meanwhile, when $\delta > 1$, the multivariate synthetic chart records a smaller value. This means, the MEWMA chart should be used to detect small and moderate mean shifts and the multivariate synthetic chart is to be used for detecting larger mean shifts.

This results tend to differ slightly from the univariate section. For the univariate, the EWMA chart is suitable for only small mean shifts whereas the synthetic chart is for moderate and larger mean shifts.

4.3.2 Comparison of the SDRL Performances of the Multivariate Synthetic and MEWMA Charts

n	**1**	**4**	**10**	**1**	**4**	**10**
δ	Multivariate Synthetic			MEWMA		
0	463.53	441.08	419.09	360.40	361.73	363.32
0.25	362.68	193.53	83.76	98.36	27.54	10.68
0.5	198.08	44.13	9.92	24.39	5.46	2.18
0.75	93.91	11.46	2.13	9.66	2.32	1.01
1	44.09	4.24	0.82	5.13	1.35	0.65
1.5	12.27	1.17	0.15	2.27	0.70	0.29
2	5.14	0.41	0.01	1.35	0.49	0.50

Table 4.4: SDRLs of the Multivariate Synthetic and MEWMA charts for n = 1, 4 and 10 based on $\text{ARL}_0 = 370$

The SDRL performances of the multivariate synthetic and the MEWMA charts are shown in the table above. By observing the table, it can be noticed that as the sample size increases, the SDRL follows a downward trend. This tells us that the variation of the run length is reduced as the sample size increases and as stated in the univariate section, this is a very good indication. The number of samples that are needed in total to detect an out-of-control signal will be much easier to estimate when there is less variation. Therefore, once again the MEWMA chart overpowers the multivariate synthetic chart. In this case, the MEWMA chart recorded smaller values throughout. But when $\delta = 2$, the SDRL values from the MEWMA chart was smaller for sample sizes, $n = 1$ and 4. For $n = 10$, the multivariate synthetic chart recorded a smaller SDRL value compared to the MEWMA chart. Therefore, it can be concluded that the multivariate synthetic chart is the preferred chart for large mean shifts and the MEWMA chart would be used for small and moderate mean shifts.

4.3.3 Comparison of the MRL Performances of the Multivariate Synthetic and MEWMA Charts

n	**1**	**4**	**10**	**1**	**4**	**10**
δ	Multivariate Synthetic			MEWMA		
0	370	370	370	370	370	370
0.25	266	141	65	109	34	16
0.5	87	16	4	33	11	6
0.75	53	7	2	17	6	4
1	31	3	1	11	5	3
1.5	12	1	1	7	3	2
2	5	1	1	5	2	2

Table 4.5: MRLs of the Multivariate Synthetic and MEWMA charts for n = 1, 4 and 10 based on $\mathrm{MRL}_0 = 370$

The MRL results that can be observed in the table above is very much similar to the ARL performance. For $0 < \delta \le 0.75$, the multivariate synthetic chart recorded much higher values compared to the MEWMA chart. Therefore, this shows that the multivariate synthetic chart will not be a suitable option for detecting small mean shifts. On the other hand, when $\delta \ge 1$, the MEWMA recorded slightly larger values which leads us to concluding that the multivariate synthetic chart should be the selected chart for detecting larger mean shifts.

4.3.4 Comparison of the ARL, MRL and SDRL Performances of the Multivariate Synthetic and MEWMA Charts (Conclusion)

To sum things up, if there are past experience that indicates that a small mean shift is expected to occur, then the MEWMA chart is to be used to monitor that specific process. On the contrary, if any collection of old data demonstrate that a moderate or a large shift usually takes place, then the multivariate synthetic chart should be the preferred option compared to the MEWMA chart.

The 3 line charts on the next few pages clearly shows the difference between the ARL, SDRL and MRL values for both the multivariate synthetic and the MEWMA charts.

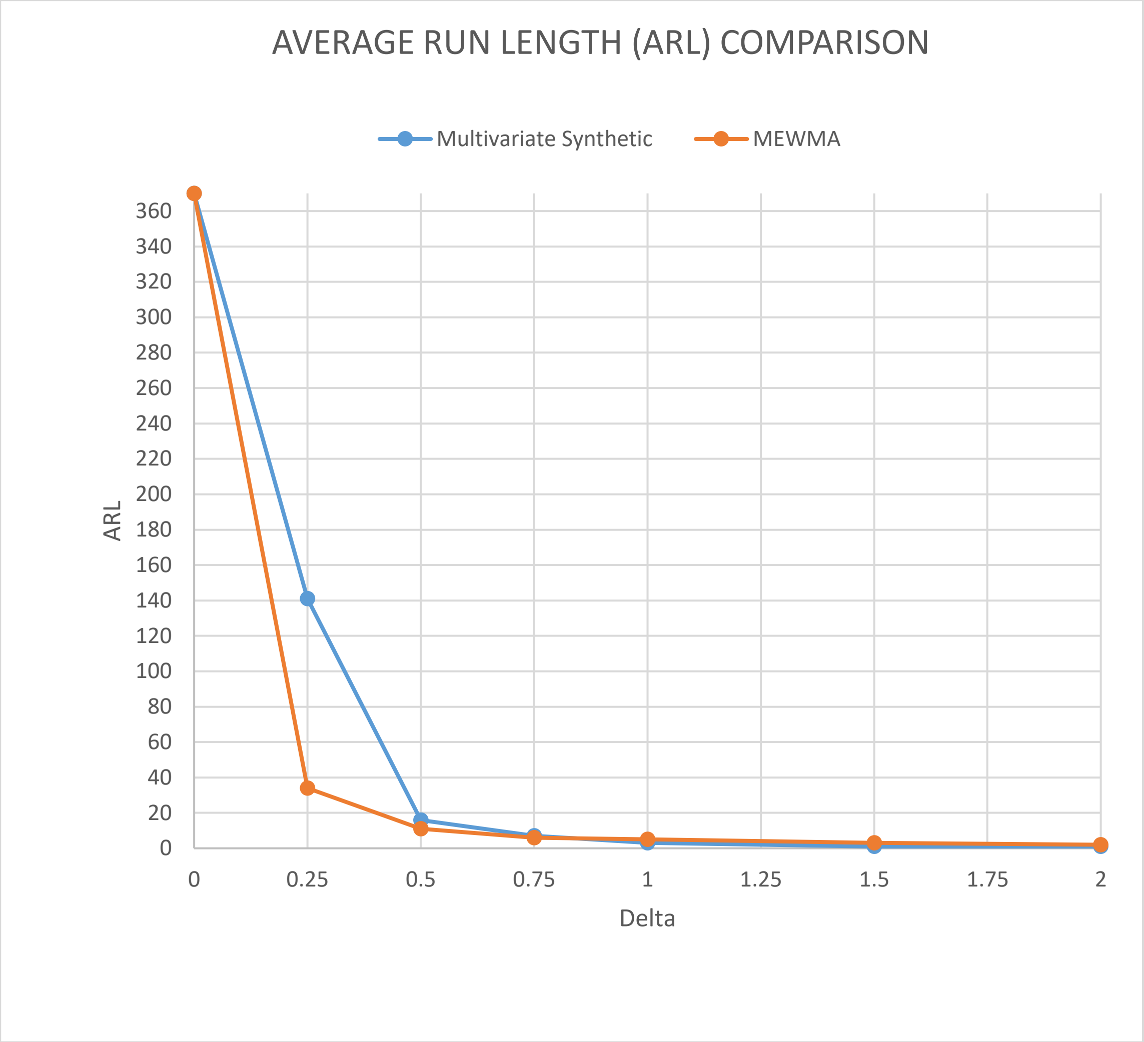

Figure 4.1: ARL comparison for the Multivariate Synthetic and MEWMA Charts

The line chart above shows that the MEWMA chart is the preferred option for small mean shifts. The line for the multivariate synthetic chart is not clearly seen for the larger mean shifts as the values of ARL are very close. However, this result can be viewed from Table 4.3.

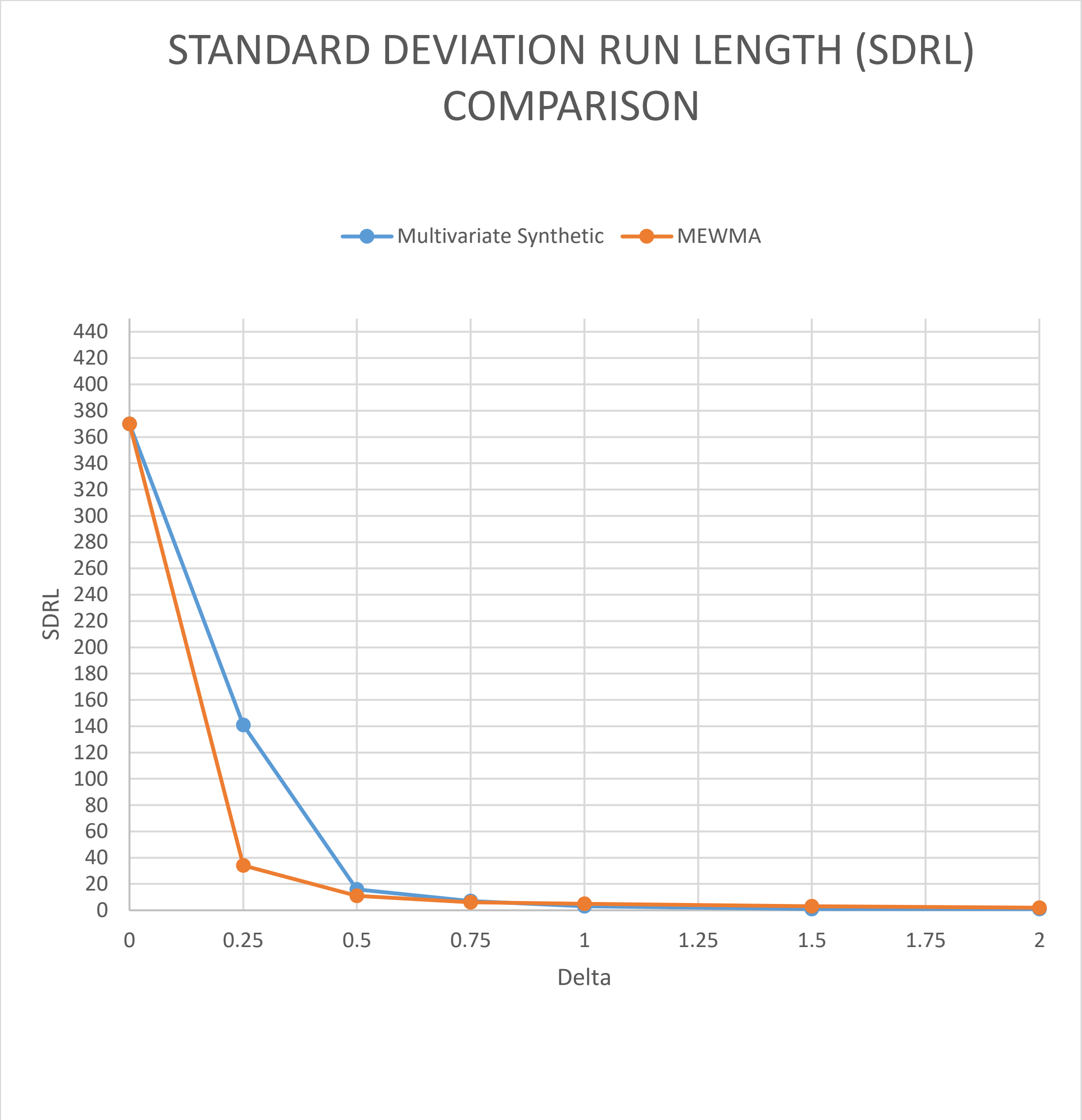

Figure 4.2: SDRL comparison for the Multivariate Synthetic and MEWMA Charts

The same observation is concluded here as well. The MEWMA chart is the most suitable chart for small mean shifts while the multivariate synthetic chart is the best option for larger mean shifts.

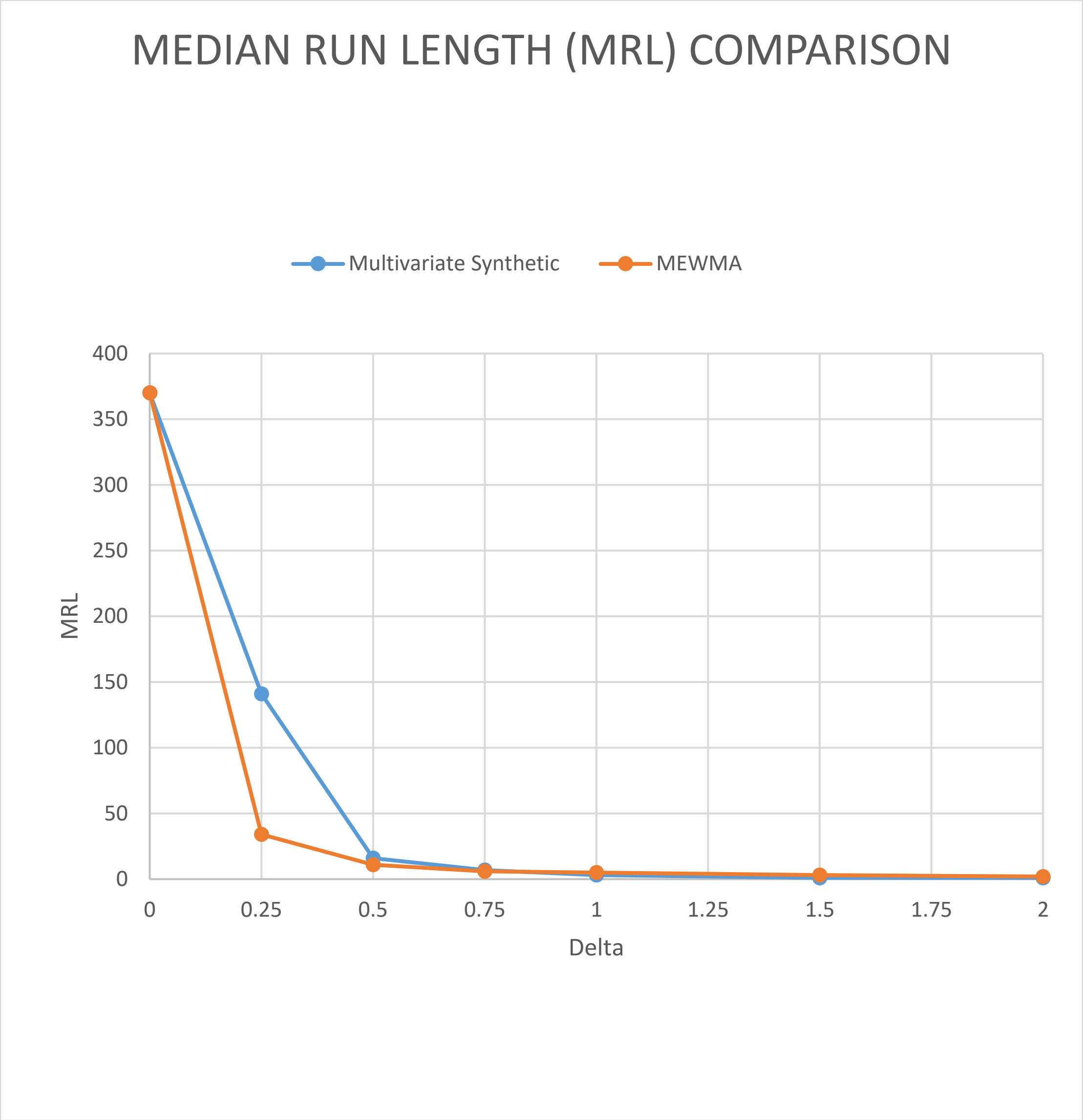

Figure 4.3: MRL comparison for the Multivariate Synthetic and MEWMA Charts

Chart 4.3 also holds the same observation and conclusion that the MEWMA is the favoured chart for small mean shifts compared to the larger ones. The multivariate synthetic chart will the better option for larger mean shifts.

CHAPTER 5

CONCLUSION

In this dissertation, the main focus was on the performances of univariate and multivariate synthetic charts. Therefore, in order to get this done, what we had done was to compute the ARLs, SDRLs and MRLs. These values were computed in both sections, the univariate and the multivariate parts. In the univariate section, the synthetic $\bar{X}$ and the EWMA charts were compared. Meanwhile, in the multivariate section, the multivariate synthetic and the MEWMA charts were compared. Comparisons were made by altering the shift in the mean, δ values and also the sample sizes. Observations were made to check and conclude which chart would be the preferred chart in detecting small, moderate and large mean shifts. In order to get these done, Mathematica programs were designed to generate values, tables were tabulated to make comparisons and last but not least, line charts were also designed to see a clearer picture of the comparisons made.

Therefore, based on all these study that was made, we have made a few conclusions. Under the univariate section, the EWMA chart is the preferred chart for small and moderate mean shifts while the synthetic chart should be used for large mean shifts. On the other hand, in the multivariate section, the same results were obtained, whereby the MEWMA chart is the chosen chart for small and moderate shifts while the multivariate synthetic chart is to be selected for large mean shifts.

All programs that were designed have been included in the Appendix section. Explanation of terms and formulas have all been done in Chapter 1 and 3.

REFERENCES

[1] Khoo, M.B.C., S. Y. The, X. Y. Chew, and W. L. Teoh (2015) 'Standard Deviation of the Run Length (SDRL) and Average Run Length (ARL) Performances of EWMA and Synthetic Charts', IACSIT International Journal of Engineering and Technology, 44(9), pp. 513-516.doi: 10.7763/IJET.2015.V7.847

[2] Khoo, M.B.C., Wong, V.H., Wu, Z. and Castagliola, P. (2012) 'Optimal design of the synthetic chart for the process mean based on median run length', *IIE Transactions*, 44(9), pp. 765–779. doi: 10.1080/0740817x.2011.609526.

[3] Wu, Z. and Spedding, T.A. (2000) A synthetic control chart for detecting small shifts in the process mean. Journal of Quality Technology, 32(1), 32-38

[4] Wu, Z., Yeo, S.H. and Spedding, T.A. (2001) A synthetic control chart for detecting fraction nonconforming increases. Journal of Quality Technology, 33(1), 104-111

[5] Gan, F.F. (1993a) An optimal design of EWMA control charts based on median run length. Journal of Statistical Computation and Simulation, 45(3), 169-184

[6] Gan, F.F (1993b) The run length distribution of a cumulative sum control chart. Journal of Quality Technology, 25(3), 205-215

[7] Montgomery, D. C., 2009. Introduction to Statistical Quality Control, 6^{th} ed. New York: Wiley & Sons

[8] Smith, Edward S. Control Charts. 1st ed. New York: N.p., 1947. Print.

[9] Besterfield, Dale H. Quality Improvement. 1st ed. Boston, Mass.:

Pearson, 2013. Print.

[10] Khoo, Michael B. C. et al. "Optimal Designs Of The Multivariate Synthetic Chart For Monitoring The Process Mean Vector Based On Median Run Length". Quality and Reliability Engineering International 27.8 (2011): 981-997. Web.

[11] Costa, A.F.B., and M.A. Rahim. "A Synthetic Control Chart For Monitoring The Process Mean And Variance". Journal of Quality in Maintenance Engineering 12.1 (2006): 81-88. Web.

[12] Haq, Abdul, and Michael B.C. Khoo. "A New Synthetic Control Chart For Monitoring Process Mean Using Auxiliary Information". Journal of Statistical Computation and Simulation 86.15 (2016): 3068-3092. Web.

APPENDICES

APPENDIX A – UNIVARIATE ANALYSIS

A.1 A Mathematica program to compute the ARL and SDRL for the synthetic chart

```
δ1 = 2;
n = 10;
L = 2;
k = 2.0845;
s = UnitVector[L + 1, 2];
Iden = IdentityMatrix[L + 1];
LL = Table[1, {i, 1, L + 1}, {j, 1, 1}];
B = 1 - N[CDF[NormalDistribution[0, 1], k - δ1 × √n]] + N[CDF[NormalDistribution[0, 1], -k - δ1 × √n]];
A = 1 - B; R = Table[0, {i, 1, L + 1}, {j, 1, L + 1}]; R[[1, 1]] = A; R[[1, 2]] = B; R[[L + 1, 1]] = A;
If[L > 1,
  Do[
   Do[
    If[j == i + 1, R[[i, j]] = A, R[[i, j]] = 0],
    {j, 1, L + 1}], {i, 2, L}]];
{ARL} = s.(Inverse[Iden - R]).LL
{SDRL} = √(2 s.(Inverse[Iden - R]).(Inverse[Iden - R]).R.LL - (ARL)^2 + ARL)
```

A.2 A Mathematica program to compute the MRL for the synthetic chart

```
δ1 = 2;
n = 10;
L = 1;
k = 2.0252;
s = UnitVector[L + 1, 2];
Iden = IdentityMatrix[L + 1];
LL = Table[1, {i, 1, L + 1}, {j, 1, 1}];
For[MRL1 = 1, B = 1 - N[CDF[NormalDistribution[0, 1], k - δ1 × √n]] + N[CDF[NormalDistribution[0, 1], -k - δ1 × √n]];
  A = 1 - B; R = Table[0, {i, 1, L + 1}, {j, 1, L + 1}]; R[[1, 1]] = A; R[[1, 2]] = B; R[[L + 1, 1]] = A;
  If[L > 1,
   Do[
    Do[
     If[j == i + 1, R[[i, j]] = A, R[[i, j]] = 0],
     {j, 1, L + 1}], {i, 2, L}]];
  {Prob} = s.(Iden - MatrixPower[R, MRL1]).LL; Prob < 0.5, MRL1 += 1];
Print[MRL1]
```

A.3 A Mathematica program to compute the ARL and SDRL for the EWMA chart

```
δ1 = 0;
n = 1;
L = 19;
k = 2.4945;
s = UnitVector[L + 1, 2];
Iden = IdentityMatrix[L + 1];
LL = Table[1, {i, 1, L + 1}, {j, 1, 1}];
B = 1 - N[CDF[NormalDistribution[0, 1], k - δ1 × √n ]] + N[CDF[NormalDistribution[0, 1], -k - δ1 × √n ]];
A = 1 - B; R = Table[0, {i, 1, L + 1}, {j, 1, L + 1}]; R[[1, 1]] = A; R[[1, 2]] = B; R[[L + 1, 1]] = A;
If[L > 1,
  Do[
   Do[
    If[j == i + 1, R[[i, j]] = A, R[[i, j]] = 0],
    {j, 1, L + 1}], {i, 2, L}]];
{ARL} = s.(Inverse[Iden - R]).LL
{SDRL} = √([2. Transpose[s].[(Inverse[Iden - R]).(Inverse[Iden - R])].R .1] - (ARL^2) + ARL)
```

A.4 A Mathematica program to compute the MRL for the EWMA chart

```
n = 1;
lambda = 0.096;
L = 0.634;
a = 0; (*change to 0, 0.25, 0.5, 0.75,1, 1.5, 2*)
m = 25; (*do not change*)
p = 2*m + 1;
mu0 = 0;
sigma0 = 1;

s = UnitVector[p, m + 1]; (*initial probability vector*)
One = Table[1, {i, 1, p}, {j, 1, 1}]; (*vector of ones*)
Iden = IdentityMatrix[p];

For[MRL1 = 1, Q = Table[0, {i, 1, p}, {j, 1, p}]; UCL = mu0 + L*sigma0; LCL = mu0 - L*sigma0; δ = (UCL - LCL)/(2*p);
  Do[Hi = LCL + (2*i - 1)*δ; Do[Hj = LCL + (2*j - 1)*δ;
    Q[[i, j]] = N[CDF[NormalDistribution[0, 1], ((Hj + δ - (1 - lambda)*Hi)/lambda - a)*Sqrt[n]]] -
      N[CDF[NormalDistribution[0, 1], ((Hj - δ - (1 - lambda)*Hi)/lambda - a)*Sqrt[n]]],
    {j, 1, p}], {i, 1, p}];
  {Prob} = s.(Iden - MatrixPower[Q, MRL1]).One; Prob < 0.5, MRL1 += 1];
Print[MRL1]
```

APPENDIX B – MULTIVARIATE ANALYSIS

B.1 A Mathematica program to compute the ARL and SDRL for the multivariate synthetic chart

```
d = 0; (*input size of mean shift - - Mahalanobis distance*)
n = 1;  (*input sample size*)
L = 61;
UCL = 9.809;
p = 2;
lambda = n × d^2; (*noncentrality parameter*)
s = UnitVector[L + 1, 2];
Iden = IdentityMatrix[L + 1];
One = Table[1, {i, 1, L + 1}, {j, 1, 1}];

B = 1 - N[CDF[NoncentralChiSquareDistribution[p, lambda], UCL]];
A = 1 - B; R = Table[0, {i, 1, L + 1}, {j, 1, L + 1}]; R[[1, 1]] = A; R[[1, 2]] = B; R[[L + 1, 1]] = A;
If[L > 1,
  Do[
   Do[
    If[j == i + 1, R[[i, j]] = A, R[[i, j]] = 0],
    {j, 1, L + 1}], {i, 2, L}]];
{ARL} = s.(Inverse[Iden - R]).One
{SDRL} = √(2. s.(Inverse[Iden - R]).(Inverse[Iden - R]).R.One - (ARL)^2 + ARL)
```

B.2 A Mathematica program to compute the MRL for the multivariate synthetic chart

```
d = 0.5; (*input size of mean shift - Mahalanobis distance*)
n = 4;  (*input sample size*)
L = 11;
UCL = 8.074;
p = 2;
lambda = n × d^2; (*noncentrality parameter*)
s = UnitVector[L + 1, 2];
Iden = IdentityMatrix[L + 1];
One = Table[1, {i, 1, L + 1}, {j, 1, 1}];

(*Computing ooc MRLδ*)
For[MRLδ = 1, B = 1 - N[CDF[NoncentralChiSquareDistribution[p, lambda], UCL]];
  A = 1 - B; R = Table[0, {i, 1, L + 1}, {j, 1, L + 1}]; R[[1, 1]] = A; R[[1, 2]] = B; R[[L + 1, 1]] = A;
  If[L > 1,
   Do[
    Do[
     If[j == i + 1, R[[i, j]] = A, R[[i, j]] = 0],
     {j, 1, L + 1}], {i, 2, L}]];
  {Prob} = s.(Iden - MatrixPower[R, MRLδ - 1]).One; Prob <= 0.5, MRLδ += 1];
MRLδ -= 1
```

B.3 A Mathematica program to compute the ARL and SDRL for the MEWMA chart

```
m1 = 25; (*input number of states for matrix H out-of-control process*)
m2 = 25; (*input number of states for matrix V out-of-control process*)

sOut = UnitVector[(2×m1 + 1) × (m2 + 1), m1 × (m2 + 1) + 1]; (*initial probability vector for ooc process*)
IdenOut = IdentityMatrix[(2×m1 + 1) × (m2 + 1)]; (*Identity matrix for ooc process*)
OneOut = Table[1, {i, 1, (2×m1 + 1) × (m2 + 1)}, {j, 1, 1}]; (*vector of ones for ooc process*)
H0 = Table[0, {i, 1, 2×m1 + 1}, {j, 1, 2×m1 + 1}];
VE = Table[0, {i, 1, m2 + 1}, {j, 1, m2 + 1}];
T = Table[0, {i, 1, (2×m1 + 1) × (m2 + 1)}, {j, 1, (2×m1 + 1) × (m2 + 1)}];
Pt = Table[0, {i, 1, (2×m1 + 1) × (m2 + 1)}, {j, 1, (2×m1 + 1) × (m2 + 1)}];

p = 2; (*input number of quality characteristics*)
δ = 0.25; (*input noncentrality parameter*)
lambda = 0.14; (*input lambda from Step 2 program*)
H = 10.568; (*input H from Step 2 program*)
n = 10; (*input sample size*)

(*Out-of-control process*)
b = (2 - lambda)/lambda; UCL = Sqrt[H]/Sqrt[b]; g1 = (2×UCL)/(2×m1 + 1); g2 = (2×UCL)/(2×m2 + 1);
Do[
  Do[ci = -UCL + (i - 0.5)×g1;
   H0[[i, j]] = N[CDF[NormalDistribution[0, 1], ((-UCL + j×g1 - (1 - lambda)×ci)/lambda - δ×Sqrt[n])]] -
     N[CDF[NormalDistribution[0, 1], ((-UCL + (j - 1)×g1 - (1 - lambda)×ci)/lambda - δ×Sqrt[n])]],
   {j, 1, 2×m1 + 1}], {i, 1, 2×m1 + 1}];

Do[
  Do[c = (((1 - lambda)×i×g2)/lambda)^2;
   If[j == 0, VE[[i + 1, j + 1]] = N[CDF[NoncentralChiSquareDistribution[p - 1, c], ((0.5)^2×g2^2)/lambda^2]],
    VE[[i + 1, j + 1]] = N[CDF[NoncentralChiSquareDistribution[p - 1, c], ((j + 0.5)^2×g2^2)/lambda^2]] -
      N[CDF[NoncentralChiSquareDistribution[p - 1, c], ((j - 0.5)^2×g2^2)/lambda^2]]],
   {j, 0, m2}], {i, 0, m2}];
P0 = KroneckerProduct[H0, VE];
```

```
c = 1;
d = 1;
Do[
  Do[
   Do[
    Do[e = (ix - (m1 + 1))^2 × g1^2 + (iy^2 × g2^2); f = (jx - (m1 + 1))^2 × g1^2 + (jy^2 × g2^2);
     If[e < UCL^2, i = 1, i = 0]; If[f < UCL^2, j = 1, j = 0];
     If[d < (2 × m1 + 1) × (m2 + 1) + 1, T[[c, d]] = i × j; Pt[[c, d]] = T[[c, d]] × P0[[c, d]]; d += 1,
      c += 1; d -= (2 × m1 + 1) × (m2 + 1); T[[c, d]] = i × j; Pt[[c, d]] = T[[c, d]] × P0[[c, d]]; d += 1]
     , {jy, 0, m2}] , {jx, 1, 2 × m1 + 1}], {iy, 0, m2}], {ix, 1, 2 × m1 + 1}];
{ARL} = sOut.(Inverse[IdenOut - Pt]).OneOut
{SDRL} = √(2. sOut.(Inverse[IdenOut - Pt]).(Inverse[IdenOut - Pt]).Pt.OneOut - ((ARL)^2) + ARL)
```

B.4 A Mathematica program to compute the MRL for the MEWMA chart

```
m1 = 25; (*input number of states for matrix H out-of-control process*)
m2 = 25; (*input number of states for matrix V out-of-control process*)

sOut = UnitVector[(2×m1 + 1) × (m2 + 1), m1 × (m2 + 1) + 1]; (*initial probability vector for ooc process*)
IdenOut = IdentityMatrix[(2×m1 + 1) × (m2 + 1)]; (*Identity matrix for ooc process*)
OneOut = Table[1, {i, 1, (2×m1 + 1) × (m2 + 1)}, {j, 1, 1}]; (*vector of ones for ooc process*)
H0 = Table[0, {i, 1, 2×m1 + 1}, {j, 1, 2×m1 + 1}];
VE = Table[0, {i, 1, m2 + 1}, {j, 1, m2 + 1}];
T = Table[0, {i, 1, (2×m1 + 1) × (m2 + 1)}, {j, 1, (2×m1 + 1) × (m2 + 1)}];
Pt = Table[0, {i, 1, (2×m1 + 1) × (m2 + 1)}, {j, 1, (2×m1 + 1) × (m2 + 1)}];

p = 2; (*input number of quality characteristics*)
δ = 0; (*input noncentrality parameter*)
lambda = 0.05; (*input lambda from Step 2 program*)
H = 8.241; (*input H from Step 2 program*)
n = 3; (*input sample size*)

(*Out-of-control process*)
```

$$b = \frac{2 - \text{lambda}}{\text{lambda}};\ UCL = \frac{\sqrt{H}}{\sqrt{b}};\ g1 = \frac{2 \times UCL}{2 \times m1 + 1};\ g2 = \frac{2 \times UCL}{2 \times m2 + 1};$$

$$\text{Do}\Big[\text{Do}\Big[ci = -UCL + (i - 0.5) \times g1;\ H0[[i, j]] = N\Big[\text{CDF}\Big[\text{NormalDistribution}[0, 1], \Big(\frac{-UCL + j \times g1 - (1 - \text{lambda}) \times ci}{\text{lambda}} - \delta \times \sqrt{n}\Big)\Big]\Big] - N\Big[\text{CDF}\Big[\text{NormalDistribution}[0, 1], \Big(\frac{-UCL + (j - 1) \times g1 - (1 - \text{lambda}) \times ci}{\text{lambda}} - \delta \times \sqrt{n}\Big)\Big]\Big], \{j, 1, 2 \times m1 + 1\}\Big], \{i, 1, 2 \times m1 + 1\}\Big];$$

$$\text{Do}\Big[\text{Do}\Big[c = \Big(\frac{(1 - \text{lambda}) \times i \times g2}{\text{lambda}}\Big)^2;\ \text{If}\Big[j == 0, VE[[i + 1, j + 1]] = N\Big[\text{CDF}\Big[\text{NoncentralChiSquareDistribution}[p - 1, c], \frac{(0.5)^2 \times g2^2}{\text{lambda}^2}\Big]\Big], VE[[i + 1, j + 1]] = N\Big[\text{CDF}\Big[\text{NoncentralChiSquareDistribution}[p - 1, c], \frac{(j + 0.5)^2 \times g2^2}{\text{lambda}^2}\Big]\Big] - N\Big[\text{CDF}\Big[\text{NoncentralChiSquareDistribution}[p - 1, c], \frac{(j - 0.5)^2 \times g2^2}{\text{lambda}^2}\Big]\Big]\Big], \{j, 0, m2\}\Big], \{i, 0, m2\}\Big];$$

```
P0 = KroneckerProduct[H0, VE];
```

```
c = 1;
d = 1;
Do[
  Do[
   Do[
    Do[e = (ix - (m1 + 1))^2 × g1^2 + (iy^2 × g2^2); f = (jx - (m1 + 1))^2 × g1^2 + (jy^2 × g2^2);
     If[e < UCL^2, i = 1, i = 0]; If[f < UCL^2, j = 1, j = 0];
     If[d < (2 × m1 + 1) × (m2 + 1) + 1, T[[c, d]] = i × j; Pt[[c, d]] = T[[c, d]] × P0[[c, d]]; d += 1,
      c += 1; d -= (2 × m1 + 1) × (m2 + 1); T[[c, d]] = i × j; Pt[[c, d]] = T[[c, d]] × P0[[c, d]]; d += 1]
     , {jy, 0, m2}] , {jx, 1, 2 × m1 + 1}], {iy, 0, m2}], {ix, 1, 2 × m1 + 1}];

(*Part 1*)
For[MRLδ = 50,
  {Prob} = sOut.(IdenOut - MatrixPower[Pt, MRLδ]).OneOut; Prob < 0.5, MRLδ += 50];
MRLδ -= 50;

(*Part 2*)
For[MRLδ = MRLδ,
  {Prob} = sOut.(IdenOut - MatrixPower[Pt, MRLδ]).OneOut; Prob < 0.5, MRLδ += 10];
MRLδ -= 10;

(*Part 3*)
For[MRLδ = MRLδ,
  {Prob} = sOut.(IdenOut - MatrixPower[Pt, MRLδ]).OneOut; Prob < 0.5, MRLδ += 1];
MRLδ
```