

Contributions to control charts for attributes data

by

Sandile Charles Shongwe

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae (Applied Statistics)

In the Department of Statistics
In the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

February 2014

Declaration

I declare that the thesis, which I hereby submit for the degree Magister Scientiae (Applied Statistics) at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: _____

Date: _____

Acknowledgements

I would like to convey my gratitude to my supervisor, Prof S. Chakraborti, and my co-supervisor, Dr M.A. Graham, for their time, guidance and constructive criticism during my study as MSc student in the Department of Statistics at the University of Pretoria. I gratefully acknowledge the comments I received from Dr S.W. Human (University of Pretoria), Prof M.M. Calzada (Loyola University of New Orleans, USA), Prof P. Castagliola (Universite de Nantes, France) and Prof M.B.C. Khoo (Universiti Sains Malaysia, Malaysia).

I would like to gratefully acknowledge the financial assistance I received from the National Research Foundation (NRF) for a Freestanding Masters Block grant No. 74392:2011, the STATOMET and the Department of Statistics at the University of Pretoria for supporting me financially, the South African Research Chairs Initiative (SARChI) for sponsoring the two SASA (South African Statistical Association) conferences I have attended and made paper presentations. Moreover, I would like to thank Mrs M. Zeelie and Mr D. du Toit for the financial assistance I received from them during the tenure of my study.

Finally, I would like to thank my family and friends for their moral support and encouragement during my study.

Summary

Chapter 1 gives a brief introduction to statistical process control (SPC) and provides definitions as well as background information regarding the research conducted in this mini-dissertation. This will aid in familiarizing the reader with concepts and terminology that are helpful to the following chapters.

We begin **Chapter 2** with a literature review of traditional methods to design Shewhart-type attributes charts and their disadvantages. It is well known that with variables data, for Case K, under the assumption of normality, a Shewhart \bar{X} chart with 3-sigma limits yields an in-control (IC) average run-length (*ARL*) equal to 370.4. However, for attributes control charts the choice of the charting constant $k = 3$ does not guarantee an IC *ARL* equal to 370.4 due to the discrete nature of the charts, as well as the fact that when the process parameters are small, the normal approximations to the binomial distribution and the Poisson distribution do not necessarily hold or hold well. In fact, attributes control charts with $k = 3$ often result in false alarm rates (*FAR*) values that are significantly different from the advertised nominal value, and this, in turn, raises questions about the efficiency of these charts. We then propose new and improved control limits for the Shewhart-type p , np , c , u charts for parameters known (Case K). It will be shown that this method yields control limits that result in IC run-length properties, such as the *FAR* and the standard deviation of the run-length (*SDRL*), that are either the same or much closer to the nominal values compared to the two traditional methods. Moreover, this method can be formulated such that it yields the same or better *ARL*-unbiased control limits compared to the traditional methods.

In **Chapter 3**, we provide a comprehensive literature review and bibliography of synthetic control charts for both univariate and multivariate cases. We consider variables (both parametric and nonparametric) control charts and attributes control charts in this review. Synthetic control charts were proposed in Wu and Spedding (2000a) and in the early 2000's there were few outputs on this topic. However, recently there is a lot of interest among researchers in this topic. Thus, there is a need for a review study, as review studies typically spark a number of new research ideas. Moreover, SAS® programs to calculate the chart parameters and the *ARL* values of the synthetic chart are given. In addition, we give a comparison study to compare the performance of the synthetic chart, the Shewhart \bar{X} chart, the 2-of-2 KL chart and the 2-of-3 KL chart.

In **Chapter 4** we illustrate that synthetic Shewhart-type attributes charts suffer from similar disadvantages as the non-synthetic counterparts discussed in Chapter 2. That is, synthetic attributes charts with $k = 3$ often result in *FAR* values that are significantly different from the advertised nominal value. Hence, we similarly propose new and improved control limits for the synthetic Shewhart-type p , np , c , u charts for parameters known (Case K). Furthermore, we show that this method yields control limits that result in IC run-length properties, such as the *FAR* and the *SDRL*, that are either the same or much closer to the nominal values compared to the two traditional methods. Moreover, this method can be formulated such that it yields the same or better *ARL*-unbiased control limits compared to the traditional methods.

Finally, **Chapter 5** wraps up this mini-dissertation with a summary of the research carried out and offers concluding remarks concerning unanswered questions and / or future research ideas.

Table of Contents

Chapter 1: Introduction	1
1.1 Statistical process control and monitoring	1
1.2 The control chart	2
1.3 Chance and Assignable causes	3
1.4 Variables and Attributes data	3
1.5 Run-length distribution	4
1.6 Types of control charts	5
1.6.1 Shewhart-type control charts	5
1.6.2 CUSUM-type control charts	5
1.6.3 EWMA-type control charts	7
1.7 Phase I and Phase II	9
1.8 Attributes control charts	10
1.9 Traditional methods to design attributes charts	11
1.9.1 k -sigma limits for the p and c charts	11
1.9.2 Conventional probability limits for the p and c charts	11
1.10 Synthetic control charts	12
1.11 Research objectives	13
1.11.1 Chapter 2	13
1.11.2 Chapter 3	14
1.11.3 Chapter 4	15
1.12 Appendix 1: Distributions considered in this study	16

Chapter 2: Modified improved probability limits (MIPL) design for the Shewhart-type attributes charts	18
2.1 Introduction	18
2.2 Properties of Shewhart-type attributes charts	19
2.3 Statistical design of the p chart	22
2.3.1 Traditional methods for designing the p chart	23
2.3.2 MIPL for the p chart	25
2.4 Illustrations and performance comparisons of the p chart methods	27
2.4.1 Example	27
2.4.2 Empirical comparison of the p chart methods	28
2.4.3 Performance	34
2.5 Statistical design of the c chart	39
2.5.1 Traditional methods for designing the c chart	40
2.5.2 MIPL for the c chart	42
2.6 Illustrations and performance comparisons of the c chart methods	43
2.6.1 Example	44
2.6.2 Empirical comparison of the c chart methods	45
2.6.3 Performance	47
2.7 Concluding remarks	52
2.8 Appendix 2: Microsoft Excel® calculations	53

Chapter 3: Synthetic quality control charts	59
An Overview	
3.1 Introduction	59
3.2 Operation and run-length characteristics of the synthetic chart for the mean	61
3.2.1 Parameters known (Case K)	62
• Zero-state analysis	
• Steady-state analysis	
3.2.2 Parameters unknown (Case U)	72
• Zero-state analysis	
• Steady-state analysis	
3.3 Univariate synthetic charts	77
3.3.1 Parametric variables charts	77
3.3.2 Nonparametric variables charts	84
3.3.3 Parametric attributes charts	86
3.4 Multivariate synthetic charts	89
3.3.1 Parametric variables charts	89
3.3.2 Nonparametric variables charts	94
3.3.3 Parametric attributes charts	95
3.5 Other synthetic-type charts	97
3.6 Concluding remarks	98
3.7 Appendices	99
3.7.1 Appendix 3A: Proofs of equations	99
3.7.2 Appendix 3B: SAS® programs	106
• Optimal search algorithm to determine k and H	
• Zero-state ARL calculation	
• Steady-state ARL calculation	
3.7.3 Appendix 3C: Comparison of four variables charts to monitor the process mean	109
3.7.4 Appendix 3D: SAS® programs for the runs-rule chart	117

Chapter 4: Modified improved probability limits (MIPL) design for the synthetic Shewhart-type attributes charts	119
4.1 Introduction	119
4.2 Properties of synthetic Shewhart-type attributes charts	121
4.3 Statistical design of the synthetic p chart	125
4.3.1 Traditional methods for designing the synthetic p chart	125
4.3.2 MIPL for the synthetic p chart	127
4.4 Illustrations and performance comparisons of the synthetic p chart methods	129
4.4.1 Example	129
4.4.2 Empirical comparison of the synthetic p chart methods	131
4.4.3 Performance	137
4.5 Statistical design of the synthetic c chart	142
4.5.1 Traditional methods for designing the synthetic c chart	142
4.5.2 MIPL for the synthetic c chart	144
4.6 Illustrations and performance comparisons of the synthetic c chart methods	146
4.6.1 Example	146
4.6.2 Empirical comparison of the synthetic c chart methods	147
4.6.3 Performance	152
4.7 Concluding remarks	156
4.8 Appendix 4: Microsoft Excel® calculations	157
Chapter 5: Summary and recommendations for future research	163
References	169

Chapter 1

Introduction

1.1 Statistical process control and monitoring

Statistical process control and monitoring is an application of a collection of statistical techniques which allows high quality products to be produced. Montgomery (2013, pp. 6-7) defines quality as fitness for use and that it is inversely proportional to variability. This means that, to improve the quality of a process or a product, we need to reduce the variability thereof. Moreover, Montgomery (2013, p. 35) states that effective quality improvement can be instrumental in increasing productivity and reducing cost. Furthermore, Montgomery (2013, p. 207) lists the following seven major statistical process control (SPC) problem solving tools that are used to assist in reducing variability and eliminating waste:

- histogram or stem-and-leaf plot,
- check sheet,
- Pareto chart,
- cause-and-effect diagram,
- defect concentration diagram,
- scatter diagram, and
- control chart.

Among the SPC tools, control charts are undeniably the most widely used for identifying changes in processes. Control charts are mainly used to distinguish between chance causes of variation and assignable causes of variation, with the Shewhart charts being the primary tools for this purpose.

1.2 The control chart

A control chart typically is a two dimensional graphic consisting of the values of a plotting (charting) statistic plotted on the vertical axis against time or subgroup number on the horizontal axis along with the associated control limits. The charting statistic and the control limits are calculated from the data which can be individual or subgroups (samples) of observations, collected sequentially over time. A typical two-sided Shewhart-type control chart (Walter A. Shewhart developed the statistical control chart concept in 1924) is shown in Figure 1.1.

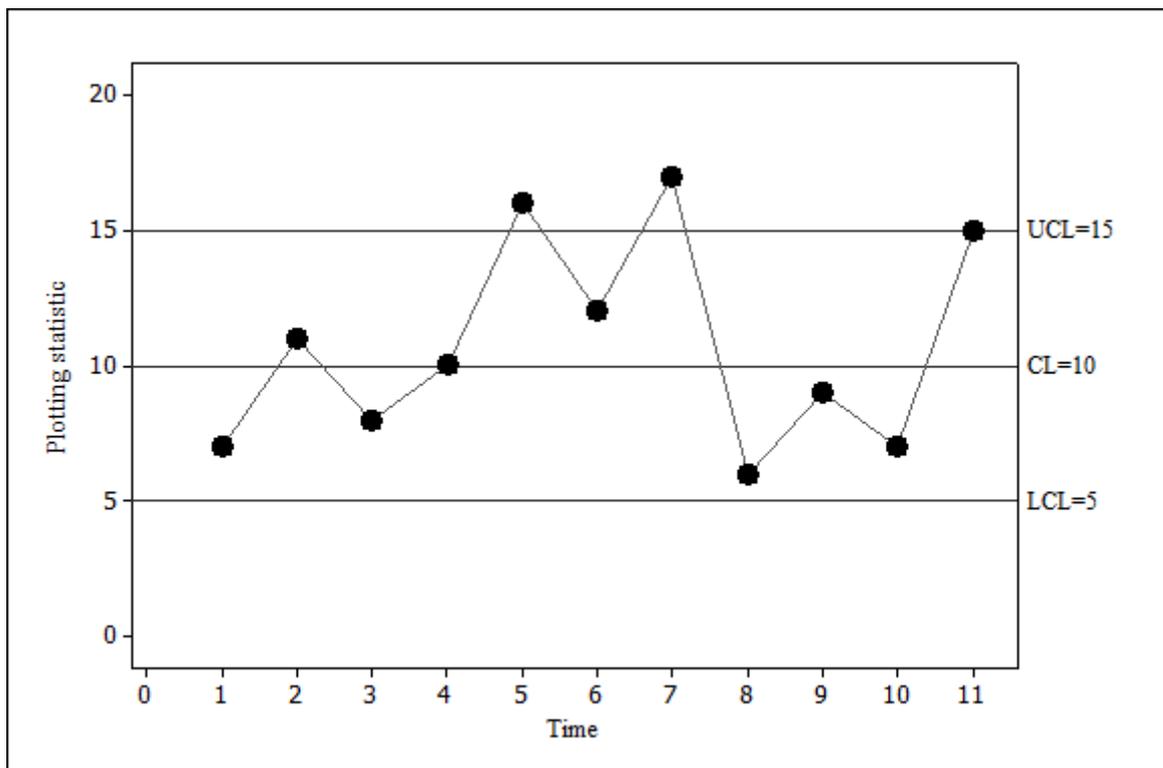


Figure 1.1. A two-sided Shewhart-type control chart

From Figure 1.1 it can be seen that a control chart usually has a center line (CL) and two horizontal lines, one on each side of the CL . The line above the CL is called the upper control limit (UCL) whereas the line below the CL is called the lower control limit (LCL). These three lines are placed on the control chart to aid the user in making an informed and objective decision whether a process is in-control (IC) or out-of-control (OOC). Note that, in some cases the UCL and LCL are not symmetric around the CL , see for example, Wu and Wang (2007). Thus, moving forward we will not mention the CL again, we only concentrate on the two important limits i.e. UCL and LCL . When a charting statistic plots on or outside either of the control limits it is said that a signal has been observed and the process is declared OOC (in Figure 1.1, a process would be thought to be OOC at times 5, 7 and 11). The corresponding event is called a signalling event. On the contrary,

when the charting statistic randomly plots between the upper and the lower control limits the process is thought to be IC and hence no signal is observed on the control chart. The corresponding event is called a non-signalling event. Montgomery (2013, p. 197) stated that control charts are popular in industries and listed five reasons for their popularity. They are:

- proven technique for improving productivity
- effective in defect prevention
- prevent unnecessary process adjustment
- provide diagnostic information
- provide information about process capability.

1.3 Chance and Assignable causes

Montgomery (2013, p. 189) stated that in any process, a certain amount of variability always exist and this natural variability is called common or chance causes. A process operating under chance cause of variation is said to be IC. However, in some cases, the source of variability is not part of the chance cause pattern. In such a situation, it is said that a process is operating in the presence of assignable causes of variation and that the process is OOC. The aim of using a control chart is to recognise and eliminate assignable causes in a process. In SPC, the pattern of chance causes is usually assumed to follow some parametric distribution (such as the normal distribution for the \bar{X} chart, see Chapter 6 in Montgomery (2013)). The charting statistic and the control limits depend on this assumption and as such the properties of these control charts are ‘exact’ only if this assumption is satisfied.

1.4 Variables and Attributes data

In statistical process control and monitoring application, data can be continuous or discrete. Quality characteristics that can be expressed in terms of a numerical measurement are called “variables” and the data collected on variables are called “variables data”, see Montgomery (2013, p. 234). Examples include dimensions such as length or width, temperature, volume etc.

However, quality characteristics that cannot be measured on a numerical scale, for example, the quality of paint on a glass container for a liquid product, are called “attributes” and the corresponding the data collected are called “attributes data” see Montgomery (2013, p. 297). To

examine attributes data, we classify them into one of the two categories called conforming and nonconforming, depending on whether the container meets the requirements on one or more quality characteristics. Examples include the number of errors or mistakes made in completing a loan application, the number of medical errors made in a hospital, etc., see Montgomery (2013, p. 297).

1.5 Run-length distribution

“The number of rational subgroups to be collected or the number of charting statistics to be plotted on a control chart before the first OOC signal is observed is the run-length of a chart”, see Human and Graham (2007). The run-length is a random variable, denoted usually by N , with a mean and variance. The most widely used chart performance metric is the mean of the run-length, referred to as the average run-length (ARL). However, since the run-length distribution is significantly right-skewed, researchers have advocated using other, more representative, measures for the assessment of chart performance. These include the standard deviation of the run-length ($SDRL$) and other percentiles of the run-length, more specifically, the median run-length (MRL), which provides additional and more meaningful information about the in-control and out-of-control performances of control charts, not given by the ARL . Some researchers such as Gan (1994), Chakraborti (2007) and Khoo et al. (2011) have advocated the use of percentiles, such as the median, for assessment of chart performance. The run-length distribution and the characteristics of the run-length distribution can be obtained using four methods, namely

- i. The exact approach (for Shewhart and some Shewhart-type charts)
- ii. The Markov chain approach
- iii. The integral equation approach
- iv. The computer simulations (the Monte Carlo) approach

For a detailed account of these methods, see Graham (2013, pp. 16 - 22). In this essay the exact approach is used which is based on a finite homogenous Markov chain; this approach is used to evaluate the run-length distribution and the characteristics of the run-length distribution of various types of control charts.

1.6 Types of control charts

There are three popular types of control charting techniques: the Shewhart chart, the cumulative sum (CUSUM) chart and the exponentially weighted moving average (EWMA) chart. Relative advantages and disadvantages of these charts are well documented in the literature. See, for example, Chapters 6, 7 and 9 in Montgomery (2013). We describe some of the charts in more detail in each of the three sections that follow.

1.6.1 Shewhart-type control charts

Shewhart-type charts are the most popular charts in practice because of their simplicity, ease of application, and the fact that these versatile charts are quite efficient in detecting moderate to large shifts. To describe the Shewhart chart in more detail, assume that $X_{i1}, X_{i2}, \dots, X_{in}$ denote a random sample (i.e. measurements on some quality characteristic) of size $n \geq 1$ from the process at time $i = 1, 2, 3, \dots$. Let W be a sample statistic that measures some quality characteristic of interest, and suppose that the mean of W is μ_W , the variance is σ_W^2 and the standard deviation of σ_W . Then the control limits and CL are given by

$$\begin{aligned}UCL &= \mu_W + k\sigma_W \\CL &= \mu_W \\LCL &= \mu_W - k\sigma_W\end{aligned}\tag{1.1}$$

where $k > 0$ is the charting constant which is a design parameter that determines the ‘distance’ of the control limits from the CL expressed in standard deviation units. When a charting statistic plots on or outside either of the control limits it is said that a signal has been observed and the process is declared OOC. Typically, a search for assignable causes is then started.

1.6.2 CUSUM-type control charts

While the Shewhart-type charts are widely known and most often used in practice because of their simplicity and global performance, other classes of charts, such as the CUSUM charts, are useful and sometimes more naturally appropriate in the process control environment in view of the sequential nature of data collection. Since the introduction of CUSUM charts by Page (1954), many researchers have examined these charts from different perspectives, see, for example, Brook and Evans (1972), Hawkins (1987, 1993) and, more recently, Abbasi et al. (2012). A comprehensive

description of the construction of CUSUM control charts is discussed in Hawkins and Olwell (1998). These charts, typically based on the cumulative sums of a charting statistic, obtained as data accumulate over time, are known to be more efficient for detecting certain types of shifts in the process; typically shifts of small magnitude. The normal theory CUSUM chart for the mean is typically based on the cumulative sum of the deviations of the individual observations (or the subgroup means) from the specified target mean.

To describe the CUSUM chart in more detail, assume that $X_{i1}, X_{i2}, X_{i3}, \dots, X_{in}$ denote a sample (subgroup) of size $n \geq 1$ on the process output at each sampling instance i for $i = 1, 2, \dots$. A statistic

$$\psi_i = \psi(X_{i1}, X_{i2}, X_{i3}, \dots, X_{in}) \quad (1.2)$$

is constructed using the data in the i^{th} sample, $i = 1, 2, \dots$. The statistic in Equation (1.2) is referred to as the basic (pivot) statistic; see Bakir (2011).

For a CUSUM chart, the deviations from the target value (say, θ_0) of the parameter are accumulated in the upward and downward directions separately, using two different statistics: one for the upward shift and the other for the downward shift.

For the upper one-sided CUSUM chart we use

$$C_i^+ = \max[0, \psi_i - k + C_{i-1}^+] \text{ for } i = 1, 2, 3, \dots \quad (1.3)$$

to detect positive deviations from the target value with starting value $C_0^+ = 0$ and the so-called reference value $k \geq 0$. A signalling event occurs for the first i such that $C_i^+ \geq h$, where $h > 0$ is the decision interval. For the lower one-sided CUSUM we use

$$C_i^- = \min[0, \psi_i + k + C_{i-1}^-] \text{ for } i = 1, 2, 3, \dots \quad (1.4)$$

or

$$C_i^{-*} = \max[0, -k - \psi_i + C_{i-1}^{-*}] \text{ for } i = 1, 2, 3, \dots \quad (1.5)$$

and is used to detect negative deviation from the target value with starting value $C_0^- = C_0^{-*} = 0$. Here a signalling event occurs for the first i such that $C_i^- \leq -h$ (if Expression (1.4) is used) or

$C_i^{-*} \geq h$ (if Expression (1.5) is used). The design parameters k and h are chosen so that the chart has a specified nominal ARL , denoted ARL_0 (or a specified nominal MRL , denoted MRL_0) and is capable of detecting a shift, specially a small shift, as soon as possible. The first step in this direction is to choose k . For the parametric CUSUM chart for the normal mean, the choice of k has been discussed by Montgomery (2013, p. 422). After choosing k , the next step is to find the decision interval h , in conjunction with the chosen k , so that a specified ARL_0 (or MRL_0) is attained. Note, however, for a discrete random variable the chances are that h cannot always be found such that the specified ARL_0 (or MRL_0) is attained exactly and hence using a conservative approach, h is found so that the attained IC ARL (or IC MRL) is less than or equal to the specified ARL_0 (or MRL_0).

1.6.3 EWMA-type control charts

The EWMA charts also take advantage of the sequentially (time ordered) accumulating nature of the data arising in a typical SPC environment and are known to be efficient in detecting smaller shifts but are easier to set up and operate than the CUSUM charts (see e.g. Montgomery (2013, p. 433). The literature on EWMA charts is enormous and continues to grow at a substantial pace (see e.g. the overview by Ruggeri et al. (2007) and the references therein). To describe the EWMA chart in more detail, assume that $X_{i1}, X_{i2}, X_{i3}, \dots, X_{in}$ denote a sample (subgroup) of size $n \geq 1$ on the process output at each sampling point i for $i = 1, 2, \dots$. The charting statistic for the EWMA control chart is defined as

$$Z_i = \lambda\psi_i + (1 - \lambda)Z_{i-1} \text{ for } i = 1, 2, 3, \dots \quad (1.6)$$

where $0 < \lambda \leq 1$ is a constant called the smoothing parameter and ψ_i is the pivot statistic defined in Equation (1.2). The starting value Z_0 is typically taken to be the target value, i.e. $Z_0 = \theta_0$. The expected value and variance of Z_i are given by

$$E(Z_i) = \theta_0 \quad (1.7)$$

and

$$VAR(Z_i) = \sigma_0^2 \left(\frac{\lambda}{2 - \lambda} \right) (1 - (1 - \lambda)^{2i}) \quad (1.8)$$

respectively, where σ_0 denotes the known process standard deviation. The exact control limits and the center line of the EWMA control chart are given by

$$UCL = E(Z_i) + L \times STDEV(Z_i) = \theta_0 + L\sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda}\right) (1 - (1-\lambda)^{2i})}$$

$$CL = E(Z_i) = \theta_0 \quad (1.9)$$

$$LCL = E(Z_i) - L \times STDEV(Z_i) = \theta - L\sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda}\right) (1 - (1-\lambda)^{2i})}$$

where $L > 0$ is a charting constant. The steady-state control limits (which are typically used when the EWMA chart has been running for several time periods so that the term $(1 - (1 - \lambda)^{2i})$ in (1.9) approaches unity) are given by

$$UCL = \theta_0 + L\sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

and

$$LCL = \theta_0 - L\sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda}\right)}. \quad (1.10)$$

The two-sided EWMA chart is constructed by plotting Z_i against the sample number i (or time). If the charting statistic Z_i falls between the two control limits, that is, $LCL < Z_i < UCL$, the process is considered to be IC. If the charting statistic Z_i falls on or outside one of the control limits, that is $Z_i \leq LCL$ or $Z_i \geq UCL$, the process is considered to be OOC and a search for assignable causes is necessary.

The two-sided EWMA can be modified to form a one-sided statistic in much the same way a CUSUM can be made into a one-sided statistic. For example, an upper one-sided EWMA is given by $Z_i^+ = \max[\theta_0, \lambda\psi_i + (1 - \lambda)Z_{i-1}]$ for $i = 1, 2, 3, \dots$ with starting value $z_0^+ = \theta_0$ where θ_0 is the IC process target value. If the charting statistic Z_i plots on or above the UCL the process is considered to be OOC and a search for assignable causes is necessary.

The design parameters L and λ are chosen so that the chart has a specified nominal ARL_0 (or MRL_0) and is capable of detecting a shift, specially a small shift, as soon as possible. Montgomery (2013, p. 436) states that “*The optimal design procedure would consist of specifying the desired in-*

control and out-of-control average run-lengths and the magnitude of the process shift that is anticipated, and then to select the combination of λ and L that provide the desired ARL performance.” The EWMA chart is designed by specifying λ and L so that a specified ARL_0 (or MRL_0) is achieved. The constant λ ($0 < \lambda \leq 1$) is the smoothing parameter ($\lambda = 1$ yields the well-known Shewhart chart) and is selected depending on the magnitude of the shift to be detected (see Table 1.1) while the constant $L > 0$ is the distance of the control limits from the CL (the larger the value of L , the wider the control limits and vice versa) and is selected in combination with the value of the smoothing parameter λ . With regard to the implementation of the EWMA chart, the first step is to choose λ . The recommendation is as follows ((see Montgomery (2013, p. 436)):

Table 1.1. Choice of the smoothing parameter λ

Magnitude of the shift of interest	Choice of λ
Small	0.05
Moderate	0.10
Large	0.20

After λ is chosen, the second step involves choosing L , so that a desired ARL_0 (or MRL_0) is attained.

1.7 Phase I and Phase II

In practice, SPC is generally divided into two phases (or stages) i.e. Phase I (also called the retrospective phase) and Phase II (also called the prospective or monitoring phase), see Montgomery (2013, p. 206). The analysis of historical or preliminary data, in order to establish that a process is IC, generally comes under what is referred to as Phase I. A process that operates at or around a desirable level or specified target with no assignable causes of variation is said to be in statistical control, or simply in-control. In Phase I, the focus is on understanding the process variability, assessing the stability of the process, investigating process improvement ideas, trying to bring the process IC by locating and eliminating any assignable causes of variability and providing estimates of the IC parameters so that effective process monitoring can begin in Phase II. A considerable amount of knowledge about a process can result from the analysis of Phase I data and control charts play a crucial role in Phase I. They help in diagnosing source(s) of assignable causes and their removal. The process of establishing control may be iterative and the control limits in this phase are usually viewed as trial limits. Once statistical control is established, the parameters are estimated and control limits are finalized based on IC data (also called reference data). Once this is ascertained, SPC moves to the next phase, called Phase II, where the control limits and / or the

estimators obtained in Phase I are used for process monitoring based on new incoming samples of data.

When the underlying parameters of the process distribution are known or specified, this is referred to as the ‘standard(s) known’ case and is denoted Case K. In contrast, if the distribution’s parameters are unknown and need to be estimated, it is typically done in Phase I, with IC data. This situation is referred to as the ‘standard(s) unknown’ case and is denoted Case U. One of the main differences between the two phases is the fact that the false alarm rate (*FAR*) (or the IC *ARL* or the IC *MRL*) is typically used to construct and evaluate Phase II control charts, whereas the false alarm probability (*FAP*) is used to construct and evaluate Phase I control charts. The *FAP* is the probability of at least one false alarm out of the comparison of all the charting statistics to the control limits simultaneously, whereas the *FAR* is the probability of a single false alarm involving only a single comparison of a charting statistic to the control limits. Various authors have studied the Phase I problem; see the review by Chakraborti et al. (2009).

1.8 Attributes control charts

Attributes control charts have been widely used to monitor discrete data in manufacturing and in service industries (i.e. non-manufacturing) processes (Montgomery (2013, p. 298)). Although an attributes chart is not as efficient as a variables control chart in finding root problems and solutions, it is an economical tool to collect and analyze some process characteristics before continuous charts can be applied (see Aebtarm and Bouguila, 2011). In addition, according to Aebtarm and Bouguila (2011), attributes control charts are more practical in many cases, for example, monitoring the number of survival patients per year is more practical than monitoring how long a patient can survive which uses a variables control chart.

Among the attributes charts, the Shewhart-type charts are the most popular. There are four major Shewhart-type attributes charts, namely the *p* chart, the *np* chart, the *c* chart and the *u* chart, see Montgomery (2013, Chapter 7) and a review by Woodall (1997). The *p* chart monitors the fraction nonconforming in a sample, whereas the *np* chart monitors the number of nonconforming items in a sample. The *c* chart monitors the number of nonconformities in a single inspection unit, whereas the *u* chart monitors the average number of nonconformities per inspection unit. Note that the designs of the latter four charts are very similar, that is, they both fit into the same statistical structure in which only the distributions assumed for each chart differ. The *p* chart and the *np* chart

are based on the normal approximation to the binomial distribution, whereas the c chart and the u chart are based on the normal approximation to the Poisson distribution.

1.9 Traditional methods to design attributes charts

In this section, we briefly summarize the traditional Shewhart-type k -sigma limits (k -SL) method and the conventional probability limits (CPL) method for Case K. To this end, first let p_0 be the known IC proportion nonconforming and c_0 be the known IC number of nonconformities in an inspection unit.

1.9.1. k -sigma limits for the p and c charts

The control limits of the p and c charts using the k -SL method are given by

$$UCL_p/LCL_p = p_0 \pm k \sqrt{\frac{p_0(1-p_0)}{n}} \quad (1.11)$$

and

$$UCL_c/LCL_c = c_0 \pm k\sqrt{c_0} \quad (1.12)$$

respectively.

1.9.2. Conventional probability limits for the p and c charts

The control limits of the p and the c charts using the CPL method are computed as follows. For the p chart, we need to find LCL_p and UCL_p such that

$$\sum_{j=0}^{\lceil nLCL_p \rceil} \binom{n}{j} p_0^j (1-p_0)^{n-j} \leq \frac{FAR_0}{2} \quad \text{and} \quad \sum_{j=\lceil nUCL_p \rceil}^n \binom{n}{j} p_0^j (1-p_0)^{n-j} \geq 1 - \frac{FAR_0}{2} \quad (1.13)$$

respectively. For the c chart, we need to find LCL_c and UCL_c such that

$$\sum_{j=0}^{\lceil LCL_c \rceil} \frac{e^{-c_0} c_0^j}{j!} \leq \frac{FAR_0}{2} \quad \text{and} \quad \sum_{j=\lceil UCL_c \rceil}^{\infty} \frac{e^{-c_0} c_0^j}{j!} \geq 1 - \frac{FAR_0}{2} \quad (1.14)$$

respectively, with $[x]$ the largest integer not exceeding x .

1.10 Synthetic control charts

Wu and Spedding (2000a) originally proposed and defined a synthetic chart as the integration of the operation of a Shewhart chart and a conforming run-length (CRL) chart. However, Scariano and Calzada (2009) proposed a generalised synthetic chart (GSC) procedure so that a synthetic chart is now defined as the integration of some control charting procedure and a CRL chart, that is, a synthetic chart consists of two “sub-charts”, one, a basic (or a standard or a classical) chart for the parameter of interest and a second, a CRL chart. Unlike the basic chart, a signal is not based on a single charting statistic falling beyond the control limits. Instead, when any sample produces a value beyond the control limits of a basic chart, then the control procedure is judged pending its effect on the CRL chart.

The conforming run-length (CRL) chart

The CRL chart was proposed by Bourke (1991) and is defined as the number of inspected units between two consecutive nonconforming units, inclusive of the nonconforming unit at the end. In most cases, we are interest in detection of process deterioration, hence the CRL chart only has a *LCL*, denoted by H . The run-length of the CRL chart follows a geometric distribution with cdf $F_{CRL}(t) = 1 - (1 - p)^t$, for $t = 1, 2, \dots$ and p is the probability that a sample is nonconforming. Consider Figure 1.2, the white and black dots denote the conforming and nonconforming units, respectively. Suppose that the process starts at $t = 0$, the CRL chart corresponding to Figure 1.1 has the following plotting/charting statistics: $CRL_1 = 5$, $CRL_2 = 2$ and $CRL_3 = 4$. The idea behind the CRL chart is that the CRL will change if the probability of obtaining a nonconforming item or probability of a plotting statistic plotting beyond the control limits increases/decreases. An OOC signal for the CRL chart is given when $CRL < H$. For further discussion on the CRL chart, see Bourke (1991).

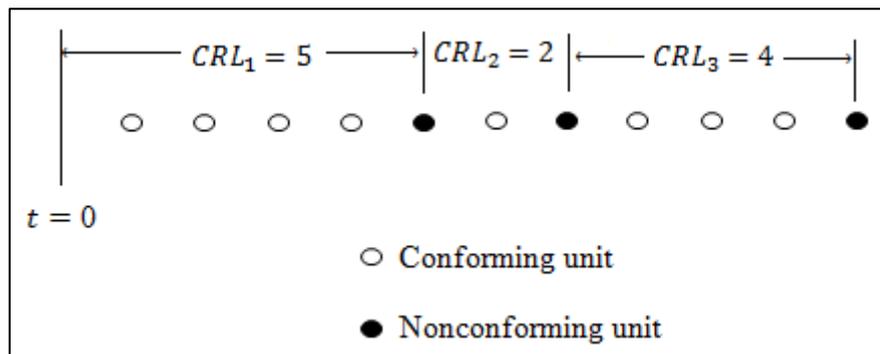


Figure 1.2: A conforming run-length chart

1.11 Research objectives

We now summarize the research questions in each of the following chapters (i.e. Chapters 2, 3, 4 and 5). In Chapters 2 and 4 we focus on proposing new and improved control limits for the Shewhart-type p , np , c , u charts and the corresponding synthetic p , np , c , u charts, respectively. Woodall and Montgomery (1999) stated that literature review papers are very important as they spark new research ideas. Motivated by this, we present a review of the literature on the synthetic control charts for univariate and multivariate data in Chapter 3. Chapter 5 provides a summary and offers some future research ideas.

1.11.1 Chapter 2

In Chapter 2 we focus on proposing new and improved control limits for the Shewhart-type p , np , c , u charts. It is well known that with variables data, for Case K, under the assumption of normality, a Shewhart \bar{X} chart with 3-sigma limits yields an IC ARL equal to 370.4. However, for attributes control charts the choice of the charting constant $k = 3$ does not guarantee an IC ARL equal to 370.4 due to the discrete nature of the charts (see Castagliola and Wu (2012)), as well as the fact that when the process parameters are small, the normal approximations to the binomial and Poisson distributions do not necessarily hold or hold well (see Montgomery (2013, p. 101)). In fact, attributes control charts with $k = 3$ often result in FAR values called attained false alarm rate ($AFAR$) values that are significantly different from the advertised nominal value, and this, in turn, raises questions about the efficiency of these charts; see Szarka and Woodall (2011).

The traditional methods are generally known to have poor IC run-length properties when the process parameters, n and/or p_0 are small. According to the recommendations in the literature, when $np_0(1 - p_0) > 5$ (i.e. when the central limit theorem approximation is good for the binomial

distribution), these methods should yield attained run-length properties that are closer to their nominal values. However, it has been shown that this is not the case; see, for example, Chakraborti and Human (2006), Wu and Wang (2007) and Castagliola and Wu (2012).

For illustration (note that more detail will be given in Chapter 2), suppose that $FAR_0 = 0.0027$, $n = 100$ and $p_0 = 0.2$ so that $np_0(1 - p_0) = 20 > 5$. Consequently, we expect that the 3-SL and CPL methods would result in an $AFAR$ close to FAR_0 . However, the 3-sigma limits method yields $AFAR = 0.00399$, which is 47.60% higher than the nominal value of 0.0027. On the other hand, the conventional probability method yields $AFAR = 0.00159$, which is 41.03% lower than the nominal value of 0.0027. It is clear for this example that these traditional control charts yield $AFAR$ values that are significantly different than the nominal values, see also Chakraborti and Human (2006), Wu and Wang (2007) and Castagliola and Wu (2012) for further confirmation. In this example, the 3-SL method chart signals more often than expected when the process is IC and the CPL method chart signals much less often than what is nominally expected, which also seems undesirable. A similar situation was observed for the c chart.

Thus, in Chapter 2, we offer a solution to this problem by considering a new and improved method of chart design to design the p , np , c , u charts for Case K. It will be shown that this method yields control limits that result in IC run-length properties, such as the $AFAR$ and the attained IC standard deviation of the run-length ($ASDRL_0$), that are either the same or much closer to the nominal values compared to the two traditional methods. Moreover, this method can be formulated such that it yields a good OOC performance.

1.11.2 Chapter 3

In Chapter 3, we provide a comprehensive literature review and bibliography of synthetic control charts for both univariate and multivariate data. We consider variables (both parametric and nonparametric) control charts and attributes control charts in this review. Ever since this chart was proposed in Wu and Spedding (2000a), there have been over 60 papers on this topic. Moreover, this review sparked a number of new research ideas and these are given in Chapter 5. In addition, a comparison study is conducted to investigate the performance of four variables control charts (synthetic chart, Shewhart \bar{X} chart, 2-of-2 KL chart and 2-of-3 KL) to monitor the process mean.

1.11.3 Chapter 4

In Chapter 4, we focus on proposing new and improved control limits for the synthetic p , np , c , u charts. It is well known that with variables data, for Case K, under the assumption of normality, a Shewhart \bar{X} chart with 3-sigma limits yields an IC ARL equal to 370.4 (see Wu and Spedding (2000a)). However, for attributes control charts the choice of the charting constant $k = 3$ does not guarantee an IC ARL equal to 370.4 due to the discrete nature of the charts (see Castagliola et al. (2013)), as well as the fact that when the process parameters are small, the normal approximations to the binomial and Poisson distributions do not necessarily hold or hold well (see Montgomery (2013, p. 101)).

The traditional methods are generally known to have poor IC run-length properties when the process parameters, n and/or p_0 are small. According to the recommendations in the literature, when $np_0(1 - p_0) > 5$ (i.e. when the central limit theorem approximation is good for the binomial distribution), these methods should yield attained run-length properties that are closer to their nominal values. However, it has been shown that this is not the case; see, see Castagliola et al. (2013).

For example (note that more detail will be given in Chapter 4), suppose that the nominal FAR (FAR_0) is equal to 0.0027, $n = 100$ and $p_0 = 0.2$ so that $np_0(1 - p_0) = 20 > 5$. Consequently, we expect that the 3-SL and CPL methods will result in $AFAR$ values close to the FAR_0 . Suppose that $H = 2$, the 3-SL method yields $AFAR = 0.00209$ which is 22.58% lower than the nominal value of 0.0027. On the other hand, the CPL method yields $AFAR = 0.00112$ which is 58.46% lower than the nominal value of 0.0027. In this example, it is clear that the traditional control charts yield $AFAR$ values that are significantly different from the FAR_0 , since for both methods, the charts will signal less often than what is nominally expected, especially for the CPL method. A similar situation was observed for the synthetic c chart.

Thus, in Chapter 4 we offer a solution to this problem by considering a new and improved method of chart design to design the synthetic p , np , c , u charts for Case K. It will be shown that this method yields control limits that result in IC run-length properties, such as the $AFAR$ and the $ASDRL_0$, that are either the same or much closer to the nominal values compared to the two traditional methods. Moreover, this method can be formulated such that it yields a good OOC performance.

1.12 Appendix 1: Distributions considered in this study

A list of the distributions considered in this study is given below along with general formulae to calculate their means and variances.

Distribution	Probability density or massfunction	Mean	Variance
Standard Normal $X \sim N(0,1)$ $x \in (-\infty, \infty)$	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$	0	1
Binomial $X \sim bi(n,p)$ $x \in \{0, 1, 2, \dots, n\}$	$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$	np	$np(1-p)$
Poisson $X \sim Poi(c)$ $x \in \{0, 1, 2, \dots\}$	$f(x) = \frac{e^{-c} c^x}{x!}$	c	c
Geometric $X \sim Poi(c)$ $x \in \{1, 2, \dots\}$	$f(x) = (1-p)^{x-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Student's t $X \sim t(v)$ $x \in (-\infty, \infty)$ $v > 0$ denotes the degrees of freedom	$f(x) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{v\pi}\Gamma(\frac{v}{2})} \left(1 + \frac{x^2}{v}\right)^{-\frac{(v+1)}{2}}$	0	$\frac{v}{v-2}$ for $v > 2$ ∞ for $v = 2$ else undefined
Gamma $X \sim GAM(\alpha, \beta)$ $x \in [0, \infty)$ $\alpha > 0$ and $\beta > 0$	$f(x) = x^{\alpha-1} \frac{e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}$	$\alpha\beta$	$\alpha\beta^2$
Logistic $X \sim Logistic(\alpha, \beta)$ $x \in (-\infty, \infty)$ $-\infty < \alpha < \infty$ and $\beta > 0$	$f(x) = \frac{e^{-(x-\alpha)/\beta}}{\beta(1+e^{-(x-\alpha)/\beta})^2}$	α	$\frac{\pi^2}{3}\beta^2$
Laplace or Double Exponential $X \sim DE(\alpha, \beta)$ $x \in (-\infty, \infty)$ $-\infty < \alpha < \infty$ and $\beta > 0$	$f(x) = \frac{1}{2\beta} e^{-\frac{ x-\alpha }{\beta}}$	α	$2\beta^2$
Uniform distribution $X \sim U(a, b)$ $x \in [a, b]$ $-\infty < a < b < \infty$	$f(x) = \frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Contaminated or Mixture Normal	Since the formulae for the Contaminated or Mixture Normal distribution is too lengthy to fit into this table, a discussion follows below.		

Contaminated or Mixture Normal distribution

The Contaminated Normal (*CN*) distribution (also referred to as the Mixture Normal distribution), is a linear combination of two normal random variables:

$$(1 - \varepsilon)N(\mu_1, \sigma_1^2) + \varepsilon N(\mu_2, \sigma_2^2),$$

where $0 < \varepsilon < 1$ denotes the level of contamination. If $X \sim (1 - \varepsilon)N(\mu_1, \sigma_1^2) + \varepsilon N(\mu_2, \sigma_2^2)$ then the pdf is given by

$$f(x) = (1 - \varepsilon)\phi(\mu_1, \sigma_1) + \varepsilon\phi(\mu_2, \sigma_2)$$

where $\phi(\mu, \sigma)$ is the pdf of a normal distribution with mean μ and variance σ^2 . The expected value and variance of the *CN* distribution are given by

$$E(X) = (1 - \varepsilon)\mu_1 + \varepsilon\mu_2$$

and

$$VAR(X) = (1 - \varepsilon)(\mu_1^2 + \sigma_1^2) + \varepsilon(\mu_2^2 + \sigma_2^2) - ((1 - \varepsilon)\mu_1 + \varepsilon\mu_2)^2,$$

respectively.

Chapter 2

Modified improved probability limits (MIPL) design for the Shewhart-type attributes charts

2.1 Introduction

Attributes control charts were introduced in Section 1.8. There are four major Shewhart-type attributes charts, namely the p chart, the np chart, the c chart and the u chart. In this chapter, we focus on the p chart and the c chart which monitors the proportion of nonconforming items in a sample and the number of nonconformities in an inspection unit, respectively. Note that when the sample size is fixed, the p chart is equivalent to the np chart (Woodall, 1997).

For a thorough account of the attributes control chart literature, see Woodall (1997). More recently, Szarka and Woodall (2011) gave a related literature review on the Bernoulli processes, with a section discussing the relevance of the p chart in a monitoring environment. Wu and Wang (2007) gave a brief discussion of the np chart using the conventional probability limits (CPL) method. Aebtarm and Bouguila (2011) presented an empirical comparison of eleven different methods to design a c control chart. Recently, Castagliola and Wu (2012) discussed the np and the c charts using the Shewhart k -sigma limits (k -SL) method when the charting parameter k is different from the typical industry value of 3. The latter authors extended on the work done in Braun (1999) and Chakraborti and Human (2006, 2008).

Shewhart control charts are generally recommended if a quick detection of a “large” shift is needed. It is well known that with variables data for Case K, under the assumption of normality, a Shewhart \bar{X} chart with 3-SL yields an IC ARL equal to 370.4. However, for attributes control charts the choice of the charting constant $k = 3$ does not guarantee an IC ARL equal to 370.4 due to the discrete nature of the charts, as well as the fact that when the process parameters are small, the normal approximation to the binomial and Poisson distributions does not

necessarily hold or hold well. In fact, attributes control charts with $k = 3$ often result in $AFAR$ values that are significantly different from the advertised nominal values, and this, in turn, raises questions about the efficiency of these charts, see Szarka and Woodall (2011).

The concept of an “improved probability limits (IPL)” chart design was proposed by Zhang et al. (2004) to design charts that have attained FAR values closer to the nominal FAR value and applied it to the geometric control chart. In this chapter we refine and modify this method and apply it to design what are called the modified improved probability limits (MIPL) p and c charts that result in better attained FAR values. In addition, these new and improved charts can be formulated such that they have the same or better OOC performance than the two traditional methods (i.e. k -SL and CPL).

The rest of the chapter is structured as follows. The run-length properties used for evaluating the statistical performance of the control charts are discussed in Section 2.2. This is followed by a discussion of the MIPL method for the p chart in Section 2.3 and the corresponding discussion for the c chart in Section 2.5; in each section, a review of the traditional methods is done. In Sections 2.4 and 2.6 examples and empirical comparisons among the three methods are done for the p and c charts, respectively, providing an insight concerning the choice of the best method. Concluding remarks are given in Section 2.7.

2.2 Properties of Shewhart-type attributes charts

Suppose that both the fraction nonconforming (p) or the average number of nonconformities (c) per inspection unit of an IC process are known (or specified), and are denoted by p_0 and c_0 , respectively. The formulas for the control limits are presented and discussed in Sections 2.3 and 2.5, respectively. Once the control limits are calculated, independent random samples (subgroups) or inspection units are typically taken at equally spaced intervals and a charting statistic, which is the proportion nonconforming (for the p chart) or the number of nonconformities (for the c chart), is plotted on the chart for the i^{th} subgroup or inspection unit, for $i = 1, 2, \dots$. If the charting statistic plots between the LCL and the UCL , the process is declared IC, otherwise, the process is declared OOC and it is said that a signal has been

observed. The run-length distribution is characterized by the probability of a no signal, denoted by β , which is a function of p and n (for the p chart) and a function of c (for the c chart). The probability of a signal is given by $1 - \beta$ and the probability of a signal when the process is IC, that is $1 - \beta_0$, is the *FAR*.

Since, in Case K, the charting statistics are iid random variables and the control chart limits are known constants, it is well known that the run-length of a Shewhart-type chart follows a geometric distribution with parameter $(1 - \beta)$. Hence the run-length pmf and cdf are given by

$$f_{RL}(r) = \beta^{r-1}(1 - \beta) \quad (2.1)$$

and

$$F_{RL}(r) = 1 - \beta^r \quad (2.2)$$

respectively, for $r = 1, 2, \dots$. The average and the standard deviation of the run-length distribution are given by

$$ARL = \frac{1}{1 - \beta} \quad (2.3)$$

and

$$SDRL = \frac{\sqrt{\beta}}{1 - \beta} \quad (2.4)$$

respectively. These quantities are used as chart performance measures.

The *FAR* is typically used to design a chart in Case K. By design one finds control limits for a given nominal *FAR* value, denoted by FAR_0 , a number such as 0.0027. Equivalently, since the *ARL* is the reciprocal of the *FAR* in Case K, one can also design the chart for a nominal IC *ARL* value such as 370.4. While this is fine for variables control charts, the actual or the attained

FAR (and the ARL) of attributes charts may not necessarily be equal to the nominal value, due to the discreteness of the distributions of the charting statistics. Thus we distinguish between the nominal and the attained values. We denote the attained FAR by $AFAR$, which is the probability that a specific control chart with a specific set of control limits (and chart design parameters), gives an OOC signal, when the process is in fact IC. That is,

$$AFAR = P(\text{Signal for a chart with a given pair of control limits} | \text{IC}) = 1 - \beta_0. \quad (2.5)$$

The problem simply is that, with the attributes charts based on normal approximation, the $AFAR$ can be substantially different from the FAR_0 . The closer the $AFAR$ value is to the FAR_0 , the better that chart is (or the chart design). The corresponding nominal ARL and $SDRL$ values are denoted by ARL_0 and $SDRL_0$, respectively, and the attained IC ARL ($= 1 / AFAR$) and the attained $SDRL$ are denoted by $AARL_0$ and $ASDRL_0$, respectively. However, when the process is OOC we denote these as $AARL_1$ and $ASDRL_1$, respectively. Other characteristics of the run-length distribution can be calculated from the pmf or the cdf of the run-length distribution.

The calculations in this chapter were done using Microsoft® Excel functions and the graphs were constructed using Minitab®. See Appendix 2A for an illustration of the calculations.

2.3 Statistical design of the p chart

Let $Y_{i,1}, \dots, Y_{i,n}$, $i = 1, 2, \dots$ and $n \geq 1$ be a sample of independent random variables $Y_i = \sum_{j=1}^n Y_{i,j} \sim \text{BIN}(p)$ where p is the proportion nonconforming units (when $p = p_0$, the process is IC). Let $nLCL_p$ and $nUCL_p$ denote the lower and upper control limits of the p chart, respectively. Montgomery (2013, p. 315) defined β , the probability of no signal (or the Type II error probability), as a function of n and p :

$$\begin{aligned}
 \beta(p, n) &= P(nLCL_p < Y_i < nUCL_p | p) && 0 < p < 1 \\
 &= P(Y_i < nUCL_p | p) - P(Y_i \leq nLCL_p | p) \\
 &= \begin{cases} \sum_{j=0}^b \binom{n}{j} p^j (1-p)^{n-j} & \text{if } nLCL_p < 0 \\ \sum_{j=0}^b \binom{n}{j} p^j (1-p)^{n-j} - \sum_{j=0}^a \binom{n}{j} p^j (1-p)^{n-j} & \text{if } nLCL_p \geq 0 \end{cases} && (2.6)
 \end{aligned}$$

The control limit constants, a and b , are related to the lower and upper control limits, respectively, and are defined in Section 2.3.1. Note that $\beta_0 = \beta(p_0, n)$, denotes the probability of no signal when the process is IC. Further, note that the probability of a signal equals

$$1 - \beta(p, n) = P(Y_i \geq nUCL_p | p) + P(Y_i \leq nLCL_p | p) \quad (2.6a)$$

and hence for a given set of control limits,

$$1 - \beta(p_0, n) = P(Y_i \geq nUCL_p | p_0) + P(Y_i \leq nLCL_p | p_0) = AFAR. \quad (2.6b)$$

2.3.1 Traditional methods for designing a p chart

In this section we briefly summarize the traditional k -SL method and the CPL method for Case K.

k -sigma limits

The control limits for the k -SL method are as given in Equation (1.11). So that, for k -SL method, the control limit constants, a and b , are given by

$$a = [nLCL_p] \text{ and } b = \begin{cases} \min\{[nUCL_p], n\} & \text{if } nUCL_p \text{ is not an integer} \\ \min\{nUCL_p - 1, n\} & \text{if } nUCL_p \text{ is an integer} \end{cases} \quad (2.7)$$

respectively, with $[x]$ the largest integer not exceeding x . These values of a and b are adopted from Chakraborti and Human (2006).

Conventional probability limits

For the CPL method, the control limits are computed as follows. For the LCL_p , using Equation (1.13) we find the largest positive integer $a = [nLCL_p]$ that makes the left tail probability, $P_i(Y_i \leq a|p_0)$, to be at most equal to $\frac{FAR_0}{2}$. Thus,

$$\sum_{j=0}^a \binom{n}{j} p_0^j (1 - p_0)^{n-j} \leq \frac{FAR_0}{2}. \quad (2.8)$$

Following this we have that $LCL_p = \frac{a}{n}$. If $a < 0$, it means that $LCL_p < 0$ and then we assume that the LCL_p does not exist since the proportion nonconforming is never negative. This arises when n and/or p_0 is small. In such a situation LCL_p is said to be not applicable (NA).

For the UCL_p , also using Equation (1.13), we find the smallest positive integer $b = [nUCL_p] - 1$ that makes the right tail probability, $P_r(Y_i \geq b + 1|p_0)$, to be at least equal to $1 - \frac{FAR_0}{2}$. Thus,

$$\sum_{j=b+1}^n \binom{n}{j} p_0^j (1 - p_0)^{n-j} \geq 1 - \frac{FAR_0}{2}. \quad (2.9)$$

Following this we have that $UCL_p = \frac{b+1}{n}$. In the event that the LCL_p is not applicable, for Equation (2.9) we find $b = nUCL_p - 1$ so that

$$P_r(Y_i \geq b + 1|p_0) \geq 1 - FAR_0. \quad (2.10)$$

Problem associated with the traditional methods

As briefly shown in Section 1.11.1, both of the traditional methods have a poor IC run-length properties (especially) when the process parameters, n and/or p_0 are small. Now, we show in detail the example discussed in that section, where $FAR_0 = 0.0027$, $n = 100$ and $p_0 = 0.2$. For the 3-SL method using Equations (1.11) and (2.7), we find $a = 8$ and $b = 31$ so that Equation (2.6b) yields $AFAR = 0.00399$, which is 47.60% higher than the nominal value of 0.0027. On the other hand, for the CPL method using Equations (2.8) and (2.9), respectively, we find $a = 8$ and $b = 33$, so that from Equation (2.6b) we obtain $AFAR = 0.00159$, which is 41.03% lower than the nominal value of 0.0027. The corresponding $AARL_0$ values are 250.63 and 628.93, respectively. It is clear for this example that these traditional control charts yield $AFAR$ and $AARL_0$ values that are significantly different than the nominal values. In this example, the 3-SL method chart signals more often than expected when the process is IC and the CPL method chart signals much less often than what is nominally expected, which also seems undesirable.

In the next section, we offer a solution to this problem by considering a new method of chart design called the modified improved probability limits (MIPL), which is an adaptation and a modification of the IPL method proposed by Zhang et al. (2004) for a geometric chart. It will be shown that this method yields control limits that result in IC run-length properties, such as the

$AFAR$ and the $ASDRL_0$, that are either the same or much closer to the nominal values compared to the two traditional methods. Moreover, the MIPL method can be formulated such that it yields a good OOC performance. Firstly we consider the p chart.

2.3.2 Modified improved probability limits (MIPL) for the p chart

In Zhang et al. (2004), the IPL method was used for the geometric control chart so that $AFAR \geq FAR_0$ (we call this an anti-conservative approach for the probability limits design). Here we modify and refine their method so that it takes into account both the conservative (i.e. $AFAR \leq FAR_0$) and the anti-conservative approach to designing probability limits. This way we obtain charts with an $AFAR$ that are much closer to the FAR_0 . Firstly we generate some set A , with all values of $a = [nLCL_p]$ satisfying some condition that will be discussed below. Then for each $a \in A$, we find the pair (a, b_1) that results in $AFAR \leq FAR_0$ and another pair (a, b_2) that results in $AFAR \geq FAR_0$. Hence, to this end, we let $S(a, b_i | n, p_0)$ for $i = 1, 2$ denote a set of control limit constants generated from set A , for some given n and p_0 . Next, let S_1 denote a subset of S with control limit constants that yield an $AFAR(a, b_1 | n, p_0) \leq FAR_0$ and similarly let S_2 denote a subset of S with control limit constants that yield an $AFAR(a, b_2 | n, p_0) \geq FAR_0$. In this case, it is easy to see that $S_1 \cup S_2 = S$. Then proceed with the following steps as in Zhang et al. (2004).

Step 1: Generate set A .

Let $a = [nLCL_p] \in A = \{\text{NA}, 0, 1, \dots, L_{max}\}$ where L_{max} is equal to the largest integer such that

$$\sum_{j=0}^{L_{max}} \binom{n}{j} p_0^j (1 - p_0)^{n-j} \leq FAR_0 \quad (2.11)$$

holds, for some specified n , FAR_0 and p_0 . Recall that “NA” stands for not applicable which implies that $LCL_p < 0$. Unlike Equation (2.8) for the CPL method, we take the left tail to be at least FAR_0 rather than $\frac{FAR_0}{2}$.

Step 2: For each $a \in A$, find the corresponding integer values of b such that

$$P_l(Y_i \leq a|p_0) + P_r(Y_i \geq b + 1|p_0) = AFAR(a, b|n, p_0) \in S \quad (2.12)$$

i.e. for each $a \in A$ in Step 1, we find $(a, b_i) \in S_i$ for $i = 1, 2$. Thus, in total we obtain $2 \times (L_{max} + 2)$ pairs of control limit constants.

Step 3: For each pair (a, b_i) in Step 2 we compute the percentage relative deviation from the FAR_0 defined by

$$D = 100 \times \left(\frac{AFAR(a, b_i|n, p_0) - FAR_0}{FAR_0} \right) \quad (2.13)$$

Step 4: Choose the pair $(a^*, b^*) \in S$ such that

$$AFAR(a^*, b^*|n, p_0) = \min_{(a, b_i) \in S} \{AFAR(a, b_i|n, p_0)\} \quad (2.14)$$

i.e. we choose the pair (a^*, b^*) that result in the minimum absolute deviation of $AFAR$ from the FAR_0 . The MIPL for the p chart are given by $LCL_p = a^*/n$ and $UCL_p = (b^* + 1)/n$.

To picture the MIPL procedure, assume that set G contains all possible pairs of control limit constants (a, b) where a and b are integers with $a < b$ and $a \in A$ (for the MIPL method), for some given n and p_0 . Some of these pairs will yield $AFAR$ values much closer to the FAR_0 and others will yield $AFAR$ values that differ significantly from the FAR_0 . Note that the control limit constants of all three methods (i.e. k -SL, CPL and MIPL) will be in set G . Furthermore, let $S \subseteq G$ as defined in Equation (2.12). That is, for each a , we find the corresponding two values of b (i.e. b_1 and b_2) such that the resulting $AFAR$ will either be at most equal or at least equal to the target FAR_0 , respectively. Then, consequently the four step procedure for the MIPL method results in the pair (a^*, b^*) , which is as close as possible to the FAR_0 , that is, a local IC optimal pair in set S and subsequently, this pair will be the global optimal pair also in set G .

Remark 1: MIPL for the np chart

Since the p chart is equivalent to the np chart when the sample size is fixed, and in this essay we assume fixed sample size (FSS) scheme, then it follows that the corresponding MIPL for the np chart are given by $LCL_{np} = a^*$ and $UCL_{np} = b^* + 1$.

In the next section, we consider a numerical example and a comparative study for the three methods considered here.

2.4 Illustrations and performance comparisons of the p chart methods

Firstly we illustrate the three methods using an example. Following this, we do an empirical comparison among the three methods by looking at different combinations of n and p_0 . Furthermore, we show that the MIPL method can be formulated such that it yields similar or better nearly ARL -unbiased control limits compared to the traditional methods.

2.4.1 Example

Example 2.1. Assume that a manufacturing production process must operate at a proportion nonconforming (fallout) of 20% and that $n = 100$ and $FAR_0 = 0.0027$.

Chart Designs

The traditional charts for this example are constructed and illustrated in Sections 1.11.1 and 2.3.1 and are displayed in Table 2.2. For the MIPL method, the calculations are as follows. From Step 1, the value of L_{max} that satisfies Equation (2.11) is equal to 9, so that $A = \{NA, 0, 1, \dots, 9\}$. For each $a \in A$, we find the corresponding b_i such that the $AFAR(a, b_i/100, 0.2)$ of these pairs is an element of S (see Step 2). Table 2.1 shows all the possible pairs of control limit constants in set S (with subsets S_1 and S_2 , as defined earlier), the $AFAR$ and the percentage relative deviation from the FAR_0 for each pair (calculated in Step 3). Then using Step 4, we see

that in Table 2.1, $(a^*, b^*) = (9, 34)$ are the optimal values for the control limits constants of the MIPL method when the process is IC, with an $AFAR$ that is only 1.12% lower than 0.0027. This is indicated by the use of grey shading in Table 2.1.

Table 2.1: All possible pairs of control limits constants in set S for the p chart using the MIPL method for $n = 100$, $p_0 = 0.2$ and $FAR_0 = 0.0027$

Set S_1				Set S_2			
$a = nLCL_p$	$b_1 = nUCL_p - 1$	$AFAR$	$ D $	$a = nLCL_p$	$b_2 = nUCL_p - 1$	$AFAR$	$ D $
NA	32	0.00155	42.58%	NA	31	0.00313	15.91%
0	32	0.00155	42.58%	0	31	0.00313	15.91%
1	32	0.00155	42.58%	1	31	0.00313	15.91%
2	32	0.00155	42.57%	2	31	0.00313	15.92%
3	32	0.00155	42.55%	3	31	0.00313	15.94%
4	32	0.00155	42.44%	4	31	0.00313	16.05%
5	32	0.00157	41.88%	5	31	0.00315	16.61%
6	32	0.00163	39.69%	6	31	0.00321	18.80%
7	32	0.00183	32.32%	7	31	0.00341	26.17%
8	32	0.00241	10.89%	8	31	0.00399	47.60%
9	34	0.00267	1.12%	9	33	0.00307	13.72%

Table 2.2 summarizes the three methods along with the $AFAR$ values, the percentage relative deviation from the FAR_0 for each pair and the $AARL_0$ and $ASDRL_0$ values, respectively.

Table 2.2: Comparison among the three methods for the p chart constants for $n = 100$, $p_0 = 0.2$ and $FAR_0 = 0.0027$

Method	(a, b)	$AFAR$	$ D $	$AARL_0$	$ASDRL_0$
3-sigma limits	(8,31)	0.00399	47.60%	250.93	250.43
Conventional probability limits	(8,33)	0.00159	41.03%	628.03	627.53
Modified improved probability limits	(9,34)	0.00267	1.12%	374.58	374.08

It is clear that, for this example, the MIPL method results in control limits with $AARL_0$ and $ASDRL_0$ values that are almost close to the target nominal value of approximately 370.

2.4.2 Empirical comparison of the p chart methods

In Example 2.1, we showed that the MIPL method yields a better $AFAR$ and a better $AARL_0$ compared to the 3-SL and CPL methods for $n = 100$ and $p_0 = 0.2$. We now investigate whether this is true for other combinations of n and p_0 values. Firstly, consider the comparison

between the 3-SL, the CPL and the MIPL methods in Figures 2.1 (a) and (b). For a fixed sample of size 50 and varying the IC proportion nonconforming from 0.01 to 0.50 in increments of 0.01, we find that the 3-SL method results in an extremely large number of false alarms when the proportion nonconforming is very small and stabilizes as the proportion nonconforming increases. For the CPL method, the resulting $FARs$ are always less than or equal to FAR_0 , however, the resulting $ASDRL_0$ is undesirable, since it is mostly much higher than the nominal value of approximately 370. Although, for small values of p_0 , the MIPL method results in high fluctuations from the nominal values, but as p_0 increases, it yields $AFAR$ and $ASDRL_0$ values that are much closer to the nominal values compared to the traditional methods.

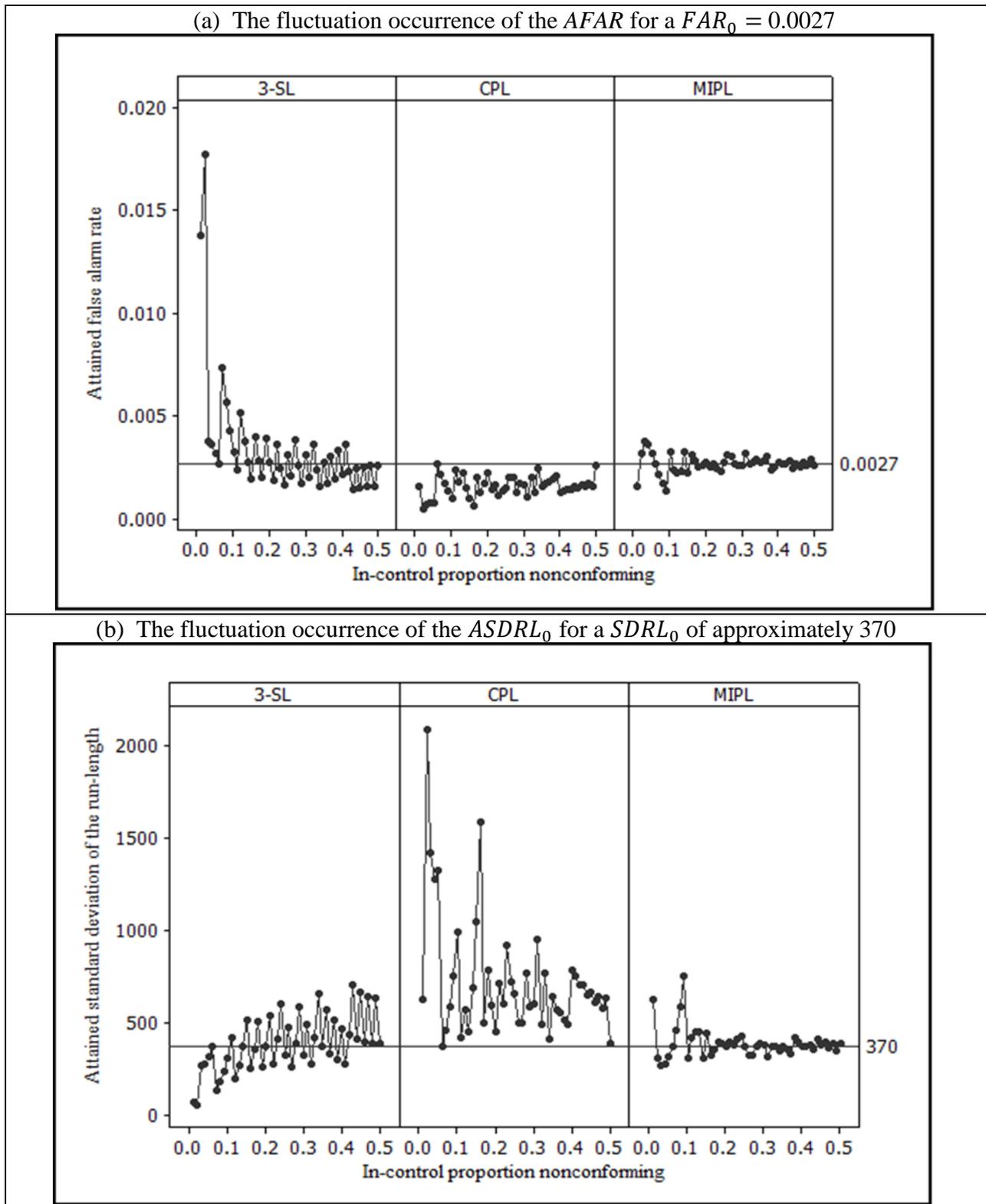


Figure 2.1: Comparison of the behavior of the run-length characteristics among the three designs of the p chart when the process is IC for $n = 50$ and varying $p_0 = 0.01(0.01)0.50$

Furthermore, for a fixed $p_0 = 0.2$ and varying the sample size, the *AFAR* behavior is shown in Figure 2.2. For the 3-SL method, there are extremely high fluctuations in the *AFAR* compared to the other methods. Although the CPL method is conservative, it is seen that it results in high $ASDRL_0$ values (see Figure 2.1 (b)). For very small sample sizes, the MIPL method has high fluctuations from the nominal value (not as bad as the 3-SL method) but once the sample size is greater than 25 the fluctuations tend to be closer to the target value of 0.0027 than the traditional methods. This pattern indicates that, as the sample size increases, the MIPL method will result in *AFAR* and $ASDRL_0$ values that are closer to target nominal values faster than the two traditional methods. Similar patterns were observed when investigating other combinations of n and p_0 .

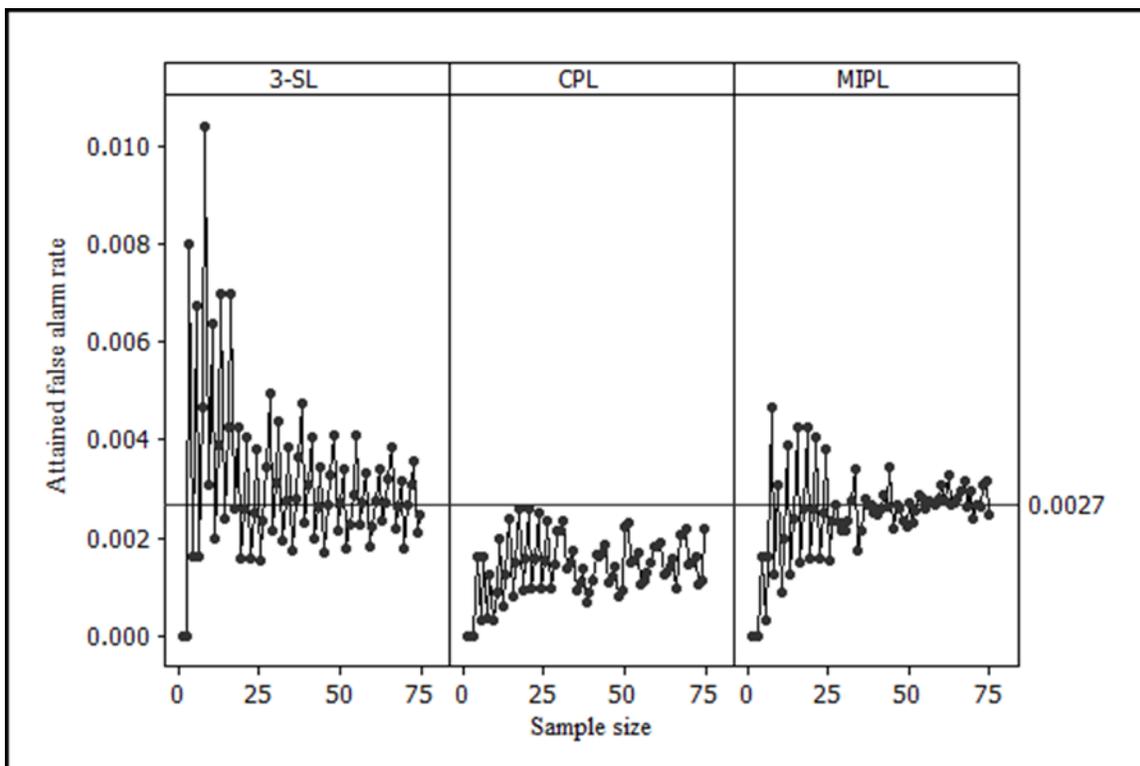


Figure 2.2: Comparison of the behavior of the *AFAR* among three designs of the p chart when the process is IC for $p_0 = 0.2$ and varying $n = 1(1)75$

In Table 2.3 we investigate more combinations of n and p_0 and we find that, for all different combinations under consideration, the MIPL method either yields the same or better $AFAR$ values with respect to the target $FAR_0 = 0.0027$. For example, for $n = 500$ and $p_0 = 0.05$, the control limits of the MIPL method yield an $AFAR = 0.00270$ whereas the 3-SL and the CPL methods yield an $AFAR$ of 0.00316 and 0.00201, respectively. However, when $n = 20$ and $p_0 = 0.2$, all three methods yield an $AFAR$ of 0.00259, in this case all three methods yield the same control limits. For $n = 5$ and $p_0 = 0.4$ or 0.5, all three methods yield an $AFAR$ of zero and consequently the $AARL_0$ and the $ASDRL_0$ approach infinity (i.e. does not exist).

Therefore, from the discussion above we see that Equation (2.14) ensures that the MIPL method yields $AFAR$ and $ASDRL_0$ values that are the same or much closer to the nominal values compared the two traditional methods.

Table 2.3: Comparison of the *AFAR* values and the corresponding percentage relative deviation (in brackets) from the nominal value of 0.0027 for the 3-SL, CPL and MIPL methods of the *p* chart in rows 1, 2 and 3, respectively, for different combinations of *n* and *p*₀

<i>n</i> / <i>p</i> ₀	0.01	0.02	0.05	0.1	0.2	0.25	0.3	0.4	0.5
5	0.04901 (1715.18%)	0.00384 (42.31%)	0.02259 (736.76%)	0.00856 (217.04%)	0.00672 (148.89%)	0.00098 (63.83%)	0.00243 (10.00%)	0 (-)	0 (-)
	0.00098 (63.70%)	0.00008 (97.13%)	0.00116 (57.11%)	0.00046 (82.96%)	0.00032 (88.15%)	0.00098 (63.83%)	0.00243 (10.00%)	0 (-)	0 (-)
	0.00098 (63.70%)	0.00384 (42.31%)	0.00116 (57.11%)	0.00046 (82.96%)	0.00032 (88.15%)	0.00098 (63.83%)	0.00243 (10.00%)	0 (-)	0 (-)
10	0.00427 (58.01%)	0.01618 (499.17%)	0.0115 (326.06%)	0.01280 (373.90%)	0.00637 (13.59%)	0.00351 (29.84%)	0.00159 (41.10%)	0.00168 (37.86%)	0.00195 (27.66%)
	0.00011 (95.78%)	0.00086 (68.00%)	0.00103 (61.91%)	0.00163 (39.45%)	0.00086 (67.99%)	0.00042 (84.60%)	0.00159 (41.10%)	0.00168 (37.86%)	0.00195 (27.66%)
	0.00427 (58.01%)	0.00086 (68.00%)	0.00103 (61.91%)	0.00163 (39.45%)	0.00086 (67.99%)	0.00351 (29.84%)	0.00159 (41.10%)	0.00168 (37.86%)	0.00195 (27.66%)
20	0.01686 (524.42%)	0.00707 (161.80%)	0.01590 (488.95%)	0.00239 (11.63%)	0.00259 (3.90%)	0.00394 (46.01%)	0.00128 (52.63%)	0.00214 (20.90%)	0.00258 (4.56%)
	0.00100 (62.83%)	0.00060 (77.79%)	0.00257 (4.67%)	0.00239 (11.63%)	0.00259 (3.90%)	0.00094 (65.36%)	0.00208 (23.08%)	0.00084 (68.85%)	0.00258 (4.56%)
	0.00100 (62.83%)	0.00060 (77.79%)	0.00257 (4.67%)	0.00239 (11.63%)	0.00259 (3.90%)	0.00394 (46.01%)	0.00208 (23.08%)	0.00214 (20.90%)	0.00258 (4.56%)
30	0.03615 (1238.81%)	0.02172 (704.36%)	0.00328 (21.57%)	0.00778 (188.28%)	0.00311 (15.22%)	0.00293 (8.45%)	0.00244 (9.74%)	0.00117 (56.68%)	0.00143 (47.00%)
	0.00022 (91.76%)	0.00030 (88.88%)	0.00057 (78.76%)	0.00202 (25.19%)	0.00214 (20.75%)	0.00100 (63.05%)	0.00094 (65.24%)	0.00117 (56.68%)	0.00143 (47.00%)
	0.00332 (22.88%)	0.00289 (7.17%)	0.00328 (21.57%)	0.00202 (25.19%)	0.00214 (20.75%)	0.00278 (3.09%)	0.00274 (1.46%)	0.00285 (5.71%)	0.00264 (2.18%)
40	0.00750 (177.68%)	0.00824 (205.22%)	0.00339 (25.63%)	0.00506 (87.52%)	0.00307 (13.67%)	0.00185 (31.39%)	0.00302 (11.75%)	0.00182 (32.73%)	0.00222 (17.72%)
	0.00069 (74.61%)	0.00118 (56.47%)	0.00071 (73.65%)	0.00147 (45.57%)	0.00112 (58.37%)	0.00159 (41.18%)	0.00145 (46.23%)	0.00182 (32.73%)	0.00222 (17.72%)
	0.00069 (74.61%)	0.00118 (56.47%)	0.00339 (25.63%)	0.00147 (45.57%)	0.00245 (9.14%)	0.00272 (0.89%)	0.00264 (2.18%)	0.00246 (9.04%)	0.00222 (17.72%)
50	0.01382 (411.75%)	0.01776 (557.71%)	0.00319 (18.09%)	0.00322 (19.26%)	0.00270 (0.14%)	0.00312 (15.41%)	0.00309 (14.53%)	0.00213 (21.07%)	0.00260 (3.62%)
	0.00160 (40.88%)	0.00048 (82.29%)	0.00076 (72.00%)	0.00100 (62.79%)	0.00222 (17.86%)	0.00151 (43.91%)	0.00166 (38.67%)	0.00128 (52.72%)	0.00260 (3.62%)
	0.00160 (40.88%)	0.00321 (18.88%)	0.00319 (18.09%)	0.00322 (19.26%)	0.00270 (0.14%)	0.00270 (0.18%)	0.00261 (3.35%)	0.00272 (0.61%)	0.00260 (3.62%)
75	0.00692 (156.30%)	0.00397 (47.09%)	0.00412 (52.53%)	0.00271 (0.51%)	0.00247 (8.49%)	0.00356 (31.93%)	0.00236 (12.73%)	0.00297 (10.06%)	0.00244 (9.47%)
	0.00097 (64.23%)	0.00077 (71.32%)	0.00123 (54.35%)	0.00138 (49.01%)	0.00217 (19.67%)	0.00116 (57.02%)	0.00142 (47.44%)	0.00201 (25.53%)	0.00244 (9.47%)
	0.00097 (64.23%)	0.00397 (47.09%)	0.00412 (52.53%)	0.00271 (0.51%)	0.00247 (8.49%)	0.00302 (11.84%)	0.00248 (8.24%)	0.00269 (0.51%)	0.00270 (0.08%)
100	0.01837 (580.52%)	0.00406 (50.45%)	0.00427 (58.30%)	0.00490 (81.57%)	0.00399 (47.60%)	0.00377 (39.76%)	0.00308 (14.20%)	0.00290 (7.04%)	0.00352 (30.28%)
	0.00053 (80.20%)	0.00093 (65.48%)	0.00146 (45.76%)	0.00113 (58.18%)	0.00159 (41.03%)	0.00171 (36.52%)	0.00205 (23.90%)	0.00207 (23.30%)	0.00179 (33.71%)
	0.00343 (27.12%)	0.00406 (50.45%)	0.00146 (45.76%)	0.00275 (1.94%)	0.00267 (1.12%)	0.00275 (1.71%)	0.00270 (0.00%)	0.00260 (3.53%)	0.00265 (1.71%)
150	0.00421 (55.97%)	0.00341 (26.19%)	0.00360 (33.45%)	0.00205 (24.11%)	0.00307 (13.70%)	0.00251 (7.01%)	0.00319 (18.16%)	0.00341 (26.22%)	0.00241 (10.91%)
	0.00085 (68.58%)	0.00095 (64.93%)	0.00100 (63.08%)	0.00145 (46.12%)	0.00203 (24.72%)	0.00173 (35.88%)	0.00232 (14.16%)	0.00196 (27.56%)	0.00241 (10.91%)
	0.00421 (55.97%)	0.00341 (26.19%)	0.00189 (29.83%)	0.00283 (4.68%)	0.00263 (2.59%)	0.00251 (7.01%)	0.00264 (2.18%)	0.00267 (1.29%)	0.00274 (1.54%)
200	0.00430 (59.09%)	0.00748 (177.01%)	0.00270 (0.01%)	0.00340 (25.96%)	0.00352 (30.28%)	0.00249 (7.86%)	0.00257 (5.00%)	0.00300 (11.03%)	0.00228 (15.58%)
	0.00101 (62.50%)	0.00253 (6.27%)	0.00156 (42.08%)	0.00127 (52.99%)	0.00189 (30.10%)	0.00182 (32.70%)	0.00194 (28.27%)	0.00237 (12.21%)	0.00228 (15.58%)
	0.00430 (59.09%)	0.00253 (6.27%)	0.00270 (0.01%)	0.00292 (8.00%)	0.00274 (1.43%)	0.00268 (0.82%)	0.00268 (0.68%)	0.00272 (0.68%)	0.00284 (5.28%)
500	0.00521 (92.89%)	0.00317 (17.31%)	0.00316 (17.07%)	0.00233 (13.86%)	0.00305 (12.79%)	0.00230 (14.73%)	0.00289 (7.05%)	0.00297 (10.14%)	0.00270 (0.12%)
	0.00065 (76.06%)	0.00107 (60.53%)	0.00201 (25.66%)	0.00215 (20.51%)	0.00207 (23.23%)	0.00230 (14.98%)	0.00244 (9.52%)	0.00220 (18.58%)	0.00270 (0.12%)
	0.00190 (29.61%)	0.00269 (0.37%)	0.00270 (0.04%)	0.00271 (0.40%)	0.00271 (0.24%)	0.00270 (0.08%)	0.00269 (0.25%)	0.00273 (1.17%)	0.00270 (0.12%)
750	0.00440 (62.91%)	0.00319 (18.33%)	0.00272 (0.82%)	0.00291 (7.88%)	0.00301 (11.36%)	0.00239 (11.44%)	0.00279 (3.38%)	0.00252 (6.75%)	0.00242 (10.47%)
	0.00127 (53.03%)	0.00154 (42.91%)	0.00189 (30.15%)	0.00185 (31.42%)	0.00220 (18.37%)	0.00239 (11.57%)	0.00244 (9.80%)	0.00252 (6.75%)	0.00242 (10.47%)
	0.00238 (11.83%)	0.00268 (0.85%)	0.00272 (0.82%)	0.00278 (3.02%)	0.00268 (0.61%)	0.00272 (0.85%)	0.00270 (0.13%)	0.00269 (0.19%)	0.00271 (0.19%)
1000	0.00333 (23.39%)	0.00266 (1.46%)	0.00305 (12.94%)	0.00270 (0.05%)	0.00303 (12.17%)	0.00243 (9.87%)	0.00267 (1.03%)	0.00267 (1.13%)	0.00265 (2.02%)
	0.00113 (58.11%)	0.00204 (24.48%)	0.00228 (15.44%)	0.00261 (3.28%)	0.00232 (14.09%)	0.00243 (9.95%)	0.00238 (12.03%)	0.00240 (11.05%)	0.00265 (2.02%)
	0.00270 (0.18%)	0.00266 (1.46%)	0.00272 (0.82%)	0.00270 (0.05%)	0.00270 (0.02%)	0.00269 (0.40%)	0.00271 (0.32%)	0.00270 (0.10%)	0.00270 (0.07%)

2.4.3 Performance

Next we investigate the performance of the p chart using these three methods when the process has a sustained shift from the IC value. It is generally known that when the process is IC, the $AARL_0$ of a control chart should be large (preferably close to the ARL_0) and when the process is OOC, the $AARL_1$ should be small. To study the OOC performance we use an ARL -curve (it shows an ARL for any possible value that parameter p can shift to (see Acosta-Mejia (1999))). For example, in Example 2.1 the control limit constants (9, 34) resulted in an $AFAR$ and an $AARL_0$ much closer to FAR_0 and ARL_0 , respectively. However, assuming that only sustained shifts with increments of 0.01 are of interest, this pair is not optimal in detecting small process deterioration, since the maximum of the ARL curve is not equal to $AARL_0$ (where $p_0 = 0.2$). That is, for some values of p ($\neq p_0$) $AARL_1 > AARL_0$. Control charts with this property were defined in Pignatiello et al. (1995) as ARL -biased charts. Acosta-Mejia (1999) showed that for the p chart, it is not always possible to obtain exact unbiasedness (because of the discrete nature of the plotting statistics).

Our aim in this section is to construct control charts that are nearly ARL -unbiased using the MIPL method, hence we proceed as follows.

Step 1: Let p^* be the value of the proportion nonconforming corresponding to the peak of the ARL curve, so that $ARL(a, b_i|n, p^*)$ is the value of the peak of the curve, with $(a, b_i) \in S$ defined in Equation (2.12). In addition, $ARL(a, b_i|n, p_0)$ is the $AARL_0$ when $p = p_0$, for some $i = 1, 2$. Then construct the ARL curve for each $(a, b_i) \in S$ for some given increment shift of size δ .

Step 2: For each pair (a, b_i) in Step 1, we compute

$$q = ARL(a, b_i|n, p^*) - ARL(a, b_i|n, p_0). \quad (2.15)$$

Note that $q = 0$ if the p chart has ARL -unbiased control limit constants.

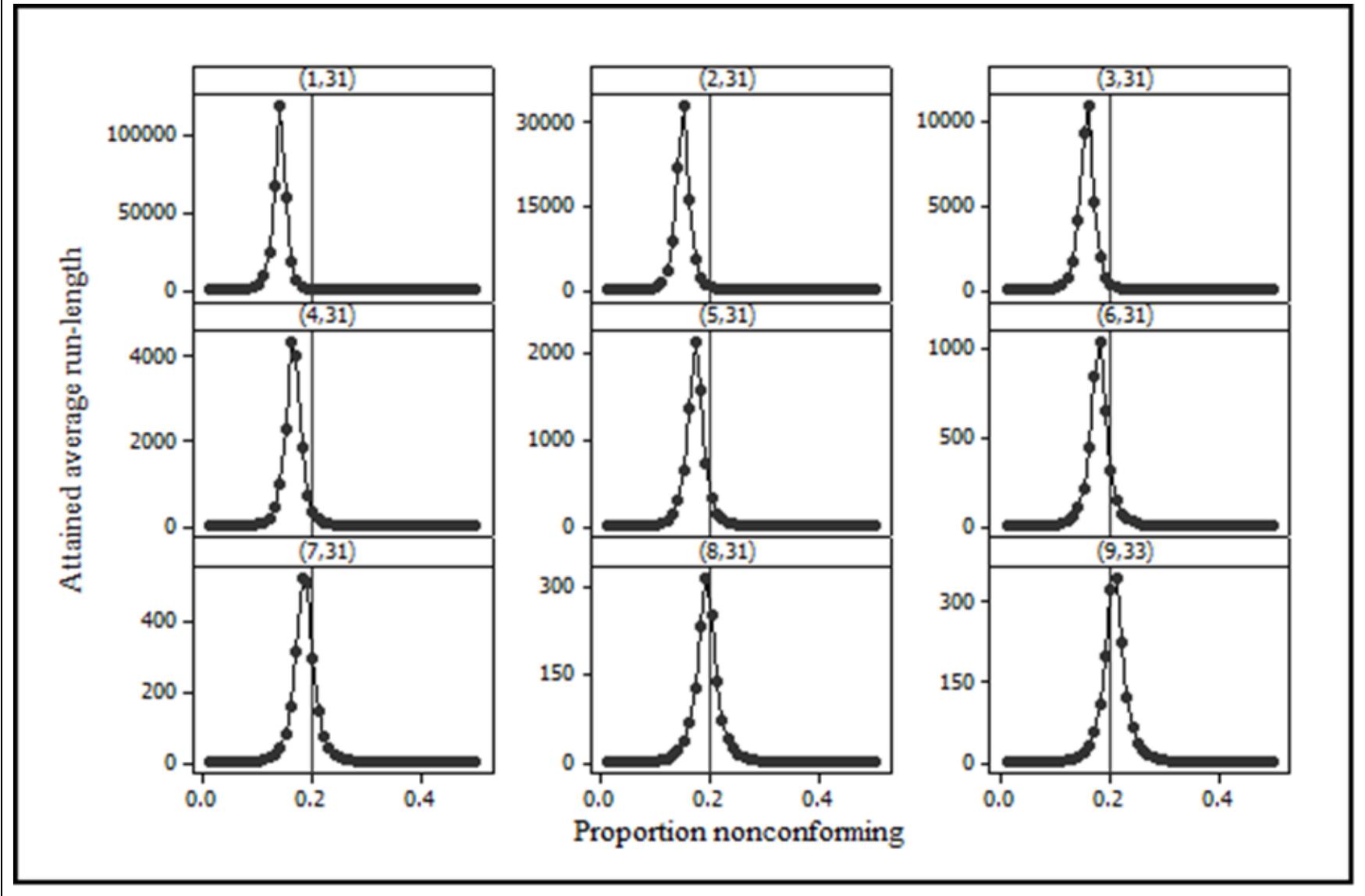
Step 3: Choose the pair $(a^\#, b^\#) \in S$ such that

$$ARL(a^\#, b^\# | n, p) = \min_{(a, b_i) \in S} ARL(a, b_i | n, p). \quad (2.16)$$

i.e. we choose the pair $(a^\#, b^\#)$ that result in the smallest value of q . Thus the nearly ARL -unbiased MIPL for the p chart are given by $LCL_p = a^\# / n$ and $UCL_p = (b^\# + 1) / n$. Note that, if there is more than one pair that satisfies Equation (2.16), then we must choose the pair that results in an $AARL_0$ closest to the ARL_0 .

To illustrate this, we use Example 2.1 to construct a p chart that results in nearly ARL -unbiased control limits. Taking the pairs (a, b_i) in S that are given in Table 2.1 as the control limit constants, we construct the ARL curves as shown in Figures 2.3 (a) and (b). Figure 2.3 (a) shows all the control limit constants in set S_1 and Figure 2.3 (b) shows all the control limit constants in set S_2 . Note that the pairs (0, 31) and (0, 32) were not plotted as the resulting ARL values (y -axis) are excessively high. It is evident that most of the control limit constants in set S have undesirable OOC values in addition to having a poor IC performance. The pair (8, 32) results in $q = 0$ (see Equation (2.15)); hence this pair results in an ARL -unbiased design for the p chart.

(a) *ARL* curves for control limit constants in set S_1 from Table 1



(b) *ARL* curves for control limit constants in set S_2 from Table 1

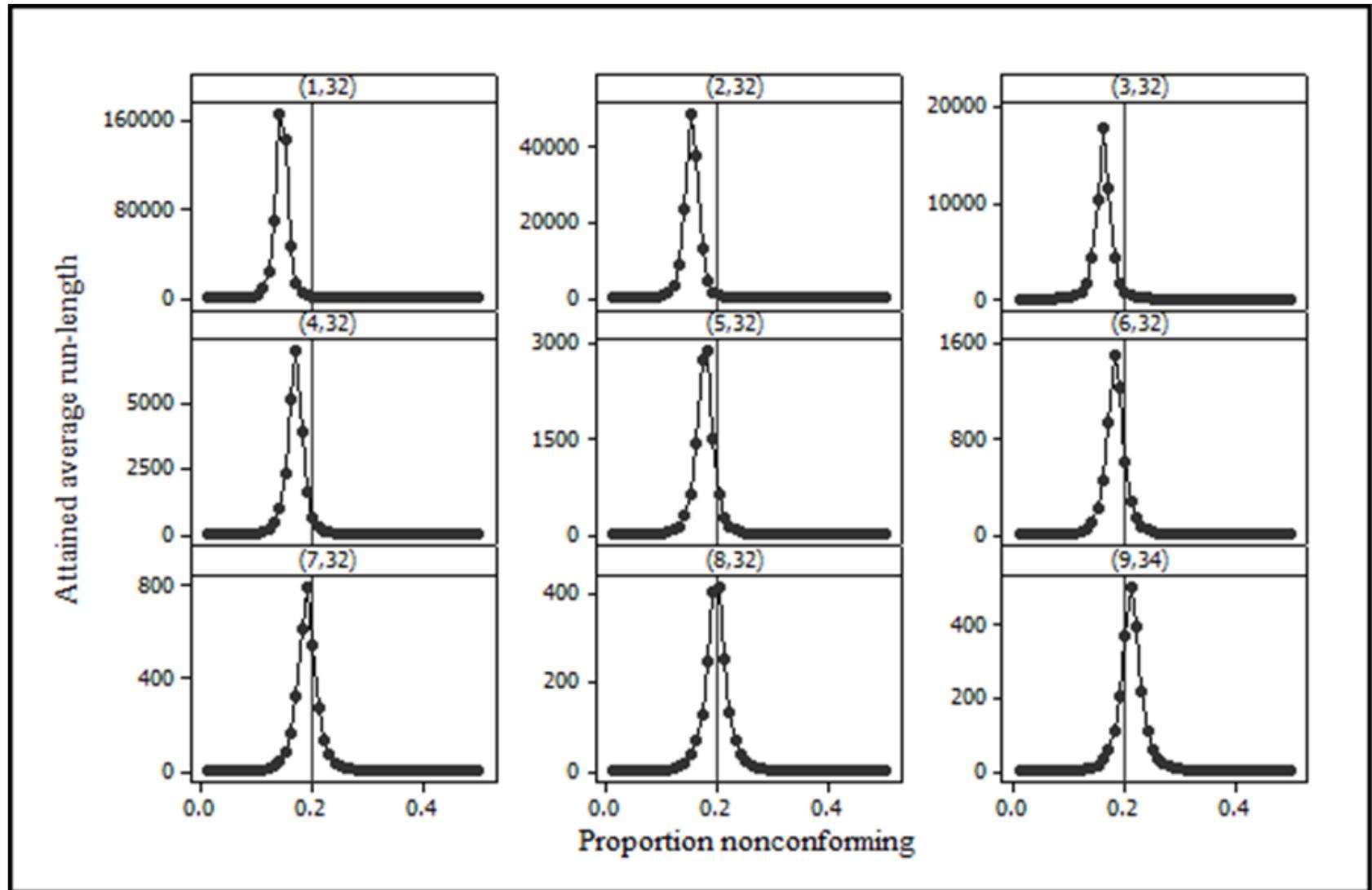


Figure 2.3: *ARL* curves of the control limit constants generated by the MIPL methods with $p_0 = 0.2$ and $n = 100$

In Figure 2.4, we plot the competing traditional methods' ARL curves along with the nearly ARL -unbiased MIPL pair (8, 32) and deduce that the 3-SL method yields ARL -biased control limits. The CPL method has ARL -unbiased control limits, but the $AARL_0 = 628.03$ is much higher than the expected value of 370.4 and, lastly, the ARL -unbiased MIPL method results in $AARL_0 = 415.66$ (much closer to 370.4 than 628.03). Therefore, if the OOC performance of the p chart is also important, the practitioner may consider taking into account Equation (2.16) rather than Equation (2.14).

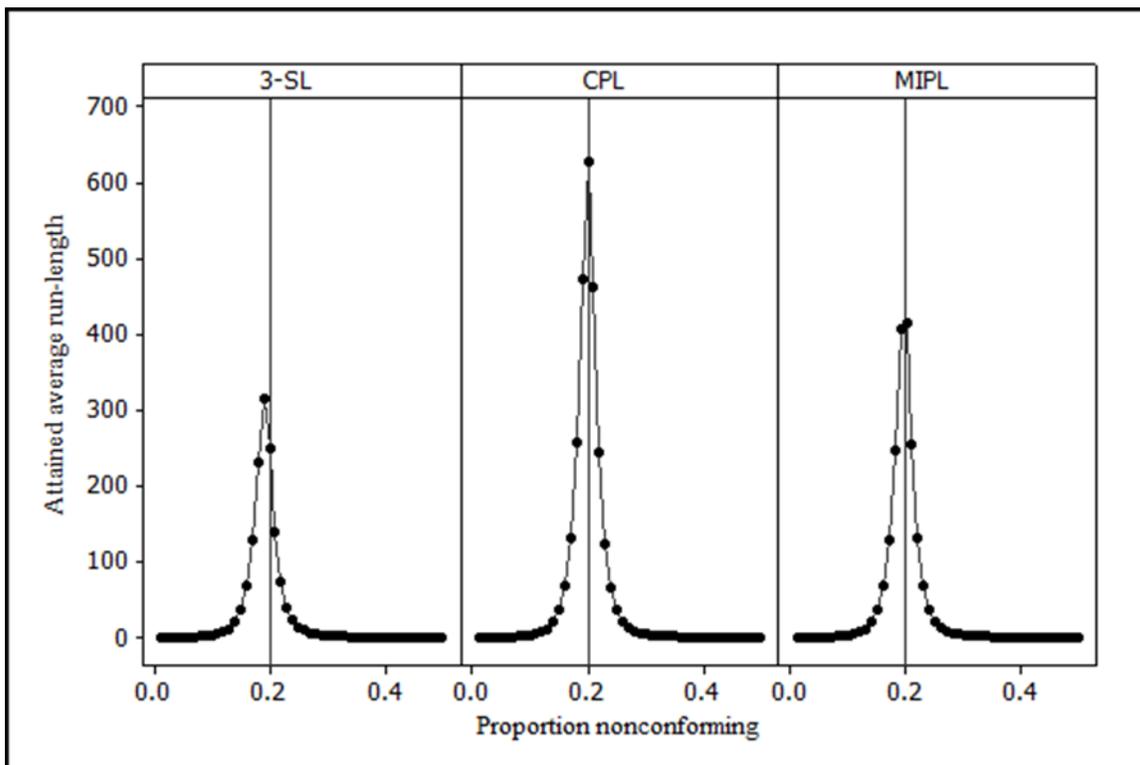


Figure 2.4: ARL curves for all three methods with $p_0 = 0.2$ and $n = 100$

Simulations indicate that when the parameters n and p are both small, the nearly ARL -unbiased MIPL method does not result in nearly ARL -unbiased control limits, however, this method never performs any worse than the 3-SL and CPL methods. In fact, for very small p_0 values, all three methods require a high value of n for the chart to be efficient, but, in most cases, the nearly ARL -unbiased MIPL method yields better performance than its competitors because the set S provides more options for the optimal pair $(a^\#, b^\#)$ compared to 3-SL and CPL methods, which have only one option for the pairs (a, b) .

Remark 2: Nearly ARL-unbiased control limits for the np chart

It follows similarly (from Remark 1) that the nearly ARL-unbiased control limits for the np chart are given by $LCL_{np} = a^\#$ and $UCL_{np} = b^\# + 1$.

2.5 Statistical design of the c chart

Let $Y_{i,1}, \dots, Y_{i,n}$, $i = 1, 2, \dots$ and $n \geq 1$ be a sample of independent random variables $Y_i = \sum_{j=1}^n Y_{i,j} \sim POI(c)$ where c is the number of nonconformities (when $c = c_0$, the process is IC). Let LCL_c and UCL_c denote the lower and upper control limits of the c chart, respectively. Montgomery (2013, p. 331) defined β , the probability of no signal (or the Type II error probability), as a function of c :

$$\begin{aligned}\beta(c) &= P(LCL_c < Y_i < UCL_c | c) && 0 < c < \infty \\ &= P(Y_i < UCL_c | c) - P(Y_i \leq LCL_c | c) \\ &= \sum_{j=0}^f \frac{e^{-c} c^j}{j!} - \sum_{j=0}^d \frac{e^{-c} c^j}{j!}.\end{aligned}\tag{2.16}$$

The control limit constants d and f are related to the lower and upper control limits, respectively, and are defined in Section 2.5.1. Note that $\beta_0 = \beta(c_0)$, denotes the probability of no signal when the process is IC. Further, note that the probability of a signal equals

$$1 - \beta(c) = P(Y_i \geq UCL_c | c) + P(Y_i \leq LCL_c | c)\tag{2.17a}$$

and

$$1 - \beta(c_0) = P(Y_i \geq UCL_c | c_0) + P(Y_i \leq LCL_c | c_0) = AFAR.\tag{2.17b}$$

2.5.1 Traditional methods for designing the c chart

k -sigma limits

The control limits for the k -SL method are given by Equation (1.12). So that, for the k -SL method, the control limit constants, d and f are given by

$$d = [LCL_c] \text{ and } f = \begin{cases} [UCL_c] & \text{if } UCL_c \text{ is not an integer} \\ UCL_c - 1 & \text{if } UCL_c \text{ is an integer} \end{cases}, \quad (2.18)$$

respectively.

Conventional probability limits

For the CPL method, the control limits are computed as follows. For the LCL_c , using Equation (1.14) we find the largest integer $d = [LCL_c]$ that makes the left tail probability, $P_l(Y_i \leq d | c_0)$, to be at most equal to $\frac{FAR_0}{2}$. Thus,

$$\sum_{j=0}^d \frac{e^{-c_0} c_0^j}{j!} \leq \frac{FAR_0}{2}. \quad (2.19)$$

Following this we have that $LCL_c = d$. If $d < 0$, it means that $LCL_c < 0$ then we assume that the LCL_c does not exist since the number of nonconformities is never negative. For Case K, this arises when c_0 is small, for instance, for the 3-SL this happens when $c_0 < 9$ (see Chakraborti and Human (2008)). In such a situation, LCL_c is said to be not applicable (NA).

For the UCL_p , again using Equation (1.14) we find the smallest integer $f = [UCL_c] - 1$ that makes the right tail probability, $P_r(Y_i \geq f + 1 | c_0)$, to be at least equal to $1 - \frac{FAR_0}{2}$, i.e.

$$\sum_{j=f+1}^{\infty} \frac{e^{-c_0} c_0^j}{j!} \geq 1 - \frac{FAR_0}{2}. \quad (2.20)$$

Following this, $UCL_c = f + 1$. In the event that LCL_c is not applicable, for Equation (2.20) we find f so that

$$P_r(Y_i \geq f + 1 | c_0) \geq 1 - FAR_0. \quad (2.21)$$

Problem associated with the traditional methods

According to the recommendations in the literature, when the normal approximation to the Poisson distribution is satisfied (i.e. central limit theorem), these methods should yield attained run-length properties that are close to the nominal values. However, it has been shown that this is not the case; see, for example, Chakraborti and Human (2008) and Castagliola and Wu (2012). For illustration, suppose that $FAR_0 = 0.0027$ and $c_0 = 20$. According to Montgomery (2013, p. 101) the normal approximation is satisfied since $c_0 \geq 15$, thus we would expect that the 3-SL and CPL methods result in an $AFAR$ much closer to the FAR_0 . However, for the 3-SL method using Equations (1.12) and (2.18), we find $d = 6$ and $f = 33$ so that Equation (2.17b) yields $AFAR = 0.00294$ which is 9.02% higher than the nominal value of 0.0027. On the other hand, for the CPL method using Equations (2.19) and (2.20) we find $a = 7$ and $b = 35$ so that Equation (2.17b) yields $AFAR = 0.00158$ which is 41.40% lower than the nominal value of 0.0027. The corresponding $AARL_0$ values are 339.72 and 632.01, respectively. It is clear for this example that these traditional control charts yield $AFAR$ and $AARL_0$ that are significantly different from the nominal values. In this example, the 3-SL method chart signals a bit more often than expected when the process is IC and the CPL method chart signals much less often than what is nominally expected, which also seems undesirable.

In the next section, we similarly offer a solution to this problem by considering a MIPL method for the c chart.

2.5.2 Modified improved probability limits (MIPL) for the c chart

Similar to the MIPL method for the p chart, we first generate set A . Further, we let $S(d, f_i|c_0)$ for $i = 1, 2$ denote a set of control limit constants generated from set A , for some given c_0 . Next, let S_1 denote a subset of S with control limits that yield an $AFAR(d, f_1|c_0) \leq FAR_0$ and similarly let S_2 denote a subset of S with control limits that yield an $AFAR(d, f_2|c_0) \geq FAR_0$. Then proceed with the following steps as in Zhang et al. (2004).

Step 1: Generate set A .

Let $d = [LCL_c] \in A = \{\text{NA}, 0, 1, 2, \dots, L_{max}\}$ where L_{max} is equal to the largest integer such that

$$\sum_{j=0}^{L_{max}} \frac{e^{-c_0} c_0^j}{j!} \leq FAR_0 \quad (2.22)$$

holds, for some FAR_0 and c_0 . “NA” stands for not applicable, it implies that $LCL_c < 0$. Note the difference between Equations (2.19) and (2.22).

Step 2: For each $d \in A$, we find the corresponding values of f such that,

$$P_l(Y_i \leq d|c_0) + P_r(Y_i \geq f + 1|c_0) = AFAR(d, f|c_0) \in S \quad (2.23)$$

i.e. for each $d \in A$ in Step 1, we find $(d, f_i) \in S_i$ for $i = 1, 2$. Thus, in total we obtain $2 \times (L_{max} + 2)$ pairs of control limit constants.

Step 3: For each pair (d, f_i) in Step 2 we compute the percentage relative deviation from the FAR_0 defined by

$$D = 100 \times \left(\frac{AFAR(d, f_i|c_0) - FAR_0}{FAR_0} \right) \quad (2.24)$$

Step 4: Choose the pair $(d^*, f^*) \in S$ such that

$$AFAR(d^*, f^* | c_0) = \min_{(d, f_i) \in S} AFAR(d, f_i | c_0). \quad (2.25)$$

i.e. we choose the pair (d^*, f^*) that result in the minimum absolute deviation of $AFAR$ from the FAR_0 . The MIPL for the c chart are given by $LCL_c = d^*$ and $UCL_c = f^* + 1$.

Similarly to the MIPL procedure of the p chart in Section 2.3.2, the four step procedure for the MIPL method of the c chart results in the pair (d^*, f^*) , which is as close as possible to the target FAR_0 , that is, a local IC optimal pair in set S and subsequently, this pair will be the global optimal pair also in set G .

Remark 3: MIPL for the u chart

Similarly, the corresponding MIPL for the u chart can be formulated by assuming that $Y_{i,j} \sim POI(u)$ where u is the OOC average number of nonconformities per inspection unit, by defining $\bar{Y}_i = \frac{1}{n} \sum_{j=1}^n Y_{i,j}$. For example, the Shewhart u chart control limits are given in Montgomery (2013, p. 324) and using u_0 instead of c_0 in Equations (2.19) and (2.20) yield the corresponding CPL method.

In the next section, we consider a numerical example and a comparative study for the three methods of the c chart considered here.

2.6 Illustration and performance comparison of the c chart methods

Firstly we illustrate the three methods using an example. Following this, an empirical comparison between the three methods is done by considering different values of the parameter c_0 . Furthermore, we show that the MIPL method can be formulated such that it yields similar or better nearly ARL -unbiased control limits compared to the traditional methods.

2.6.1 Example

Example 2.2. Assume that a manufacturing process produces inspection units that normally have 20 defects with $FAR_0 = 0.0027$.

Chart Designs

The traditional charts for this example have been found in Section 2.5.1 and are displayed in Table 2.5. For the MIPL method, the calculations are as follows. From Step 1, the value of L_{max} that satisfies Equation (2.22) is 8, so that $A = \{NA, 0, 1, \dots, 8\}$. For each $d \in A$, we find the corresponding f_i such that $AFAR(d, f_i|20)$ of these pairs is an element of S (see Step 2). Table 2.4 shows all possible pairs of control limit constants in set S (with subsets S_1 and S_2), the $AFAR$ and the percentage relative deviation from the FAR_0 for each pair (calculated in Step 3). Then using Step 4, we see that in Table 2.4, the pair $(d^*, f^*) = (4, 33)$ are the optimal pair of control limits of the MIPL method when the process is IC, with an $AFAR$ that is 0.20% higher than the 0.0027. This is indicated by the use of grey shading in Table 2.4.

Table 2.4: All possible pairs of control limit constants in set S for the c chart using the MIPL method

Set S_1				Set S_2			
$d = LCL_c$	$f = UCL_c - 1$	$AFAR$	$ D $	$d = LCL_c$	$f = UCL_c - 1$	$AFAR$	$ D $
NA	33	0.00269	0.43%	NA	32	0.00473	75.09%
0	33	0.00269	0.43%	0	32	0.00473	75.09%
1	33	0.00269	0.43%	1	32	0.00473	75.09%
2	33	0.00269	0.41%	2	32	0.00473	75.11%
3	33	0.00269	0.31%	3	32	0.00473	75.21%
4	34	0.00151	44.22%	4	33	0.00271	0.20%
5	34	0.00156	42.19%	5	33	0.00276	2.24%
6	34	0.00174	35.40%	6	33	0.00294	9.02%
7	34	0.00227	16.01%	7	33	0.00347	28.41%
8	36	0.00251	7.03%	8	35	0.00289	7.07%

Table 2.5 summarizes the three charting methods (control limits) along with the $AFAR$ and the percentage relative deviation from the FAR_0 for each pair, $AARL_0$ and $ASDRL_0$ values, respectively.

Table 2.5: Comparison among the three methods of the c chart for a $FAR_0 = 0.0027$ and $c_0 = 20$

Method	(d, f)	$AFAR$	$ D $	$AARL_0$	$ASDRL_0$
3-sigma limits	(6,33)	0.00294	9.02%	339.72	339.22
Conventional probability limits	(7,35)	0.00158	41.40%	632.01	631.51
Modified improved probability limits	(4,33)	0.00271	0.20%	369.63	369.13

It is clear that, for this example, the MIPL method results in control limit constants with much improved $AARL_0$ and $ASDRL_0$ values compared to the traditional 3-SL and CPL methods. For this example, $c_0 = 20$, the control limits $(d, f) = (4, 33)$ are the only values that ensure that we get as close as possible to the nominal ARL and $SDRL$ values.

2.6.2 Empirical comparison of the c chart methods

In Example 2.2, we showed that the MIPL method yields better $AFAR$ and $AARL_0$ values when compared to the 3-SL and CPL methods for $c_0 = 20$. We now investigate whether this is true for other values of c_0 . In Figures 2.5 (a) and (b), the $AFAR$ and $ASDRL_0$ values for small c_0 are not close to their respective nominal values, that is, when the normal approximation to the Poisson distribution is not satisfied the performance of the c chart is severely degraded (especially for 3-SL). However, as the process parameter increases, the $AFAR$ and $ASDRL_0$ values fluctuate more or less around the nominal values for the MIPL method. Whereas, the 3-SL method results in relatively higher false alarms for most of the process parameters and the CPL method has very high values for the $ASDRL_0$ (which is undesirable when the process is IC). Thus, in Figure 2.5, we see that the MIPL approach would be a preferred method to design the c chart when the process is IC.

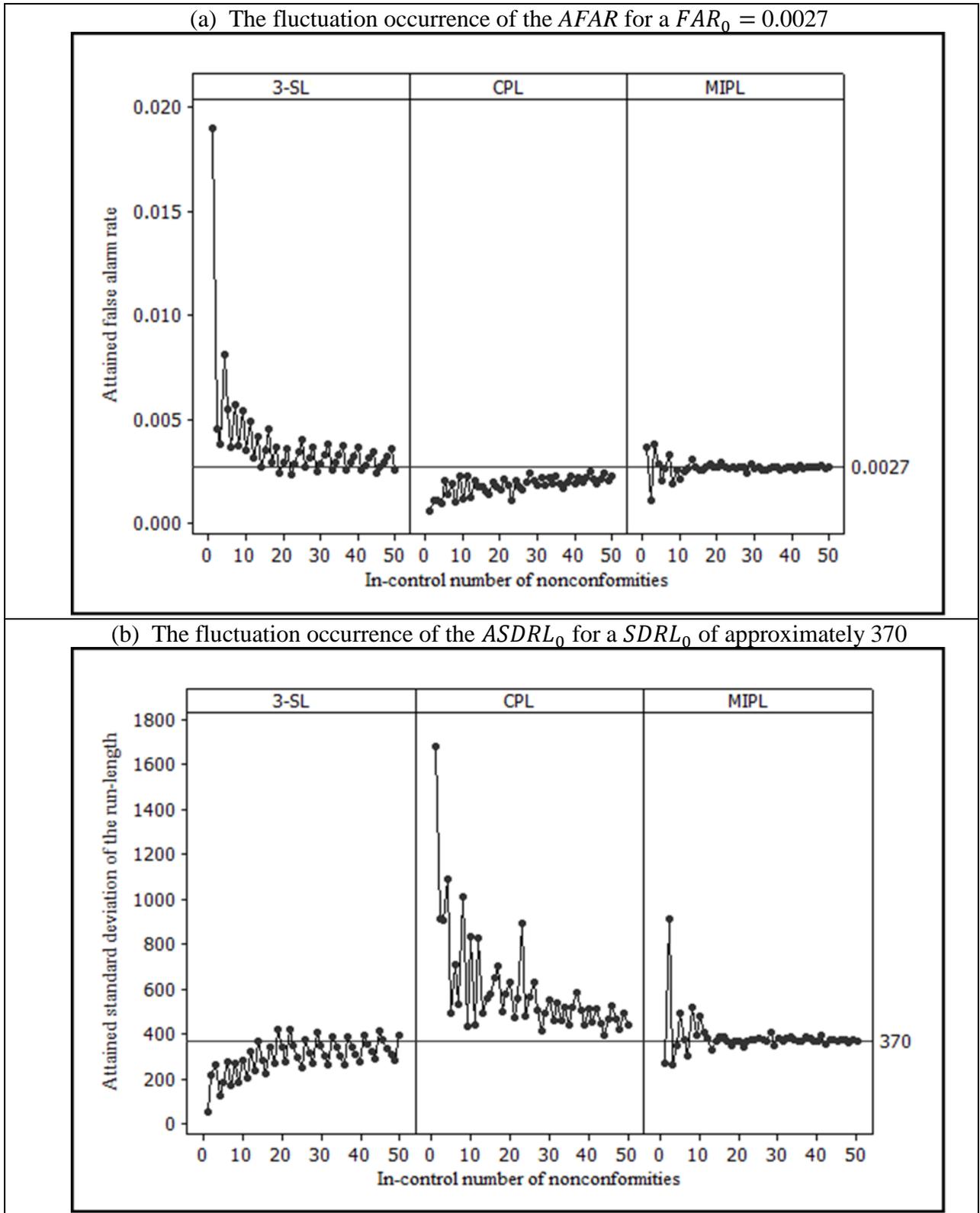


Figure 2.5: Comparison of the behavior of the run-length characteristics among three methods of the c chart when process is IC

2.6.2 Performance

Next we investigate the performance of the c chart using these three methods when the process has a sustained shift from the IC value. For example, in Example 2.2, the pair (4, 33) resulted in $AFAR$ and $AARL_0$ values much closer to FAR_0 and ARL_0 for the MIPL method, however, assuming that only sustained shifts with increments of 1 are of interest, this pair will not be optimal in detecting process improvement, since the maximum of the ARL curve for this pair is not equal to the $AARL_0$ (where $c_0 = 20$), see Figure 2.6 (b). From Figures 2.6 (a) and (b) it can be seen that there is a pair such that for all shifts, $AARL_1 < AARL_0$. For the c chart, it is not always possible to obtain exact unbiasedness (because of the discrete nature of the Poisson distribution).

Thus, if the objective is to construct a c chart such that it has nearly ARL -unbiased control limits, we need to proceed as follows.

Step 1: Let c^* be the value of the proportion nonconforming corresponding to the peak of the ARL curve, so that $ARL(d, f_i | c^*)$ is the value of the peak of the curve, with $(d, f_i) \in S$ defined in Equation (2.23). In addition, $ARL(d, f_i | c_0)$ is the $AARL_0$ when $c = c_0$, for $i = 1, 2$. Then construct the ARL curve for each $(d, f_i) \in S$ for some given increment shift of size δ .

Step 2: For each pair (d, f_i) in Step 1, we compute

$$q = ARL(d, f_i | c^*) - ARL(d, f_i | c_0). \quad (2.26)$$

Note that $q = 0$ if the c chart has ARL -unbiased control limit constants.

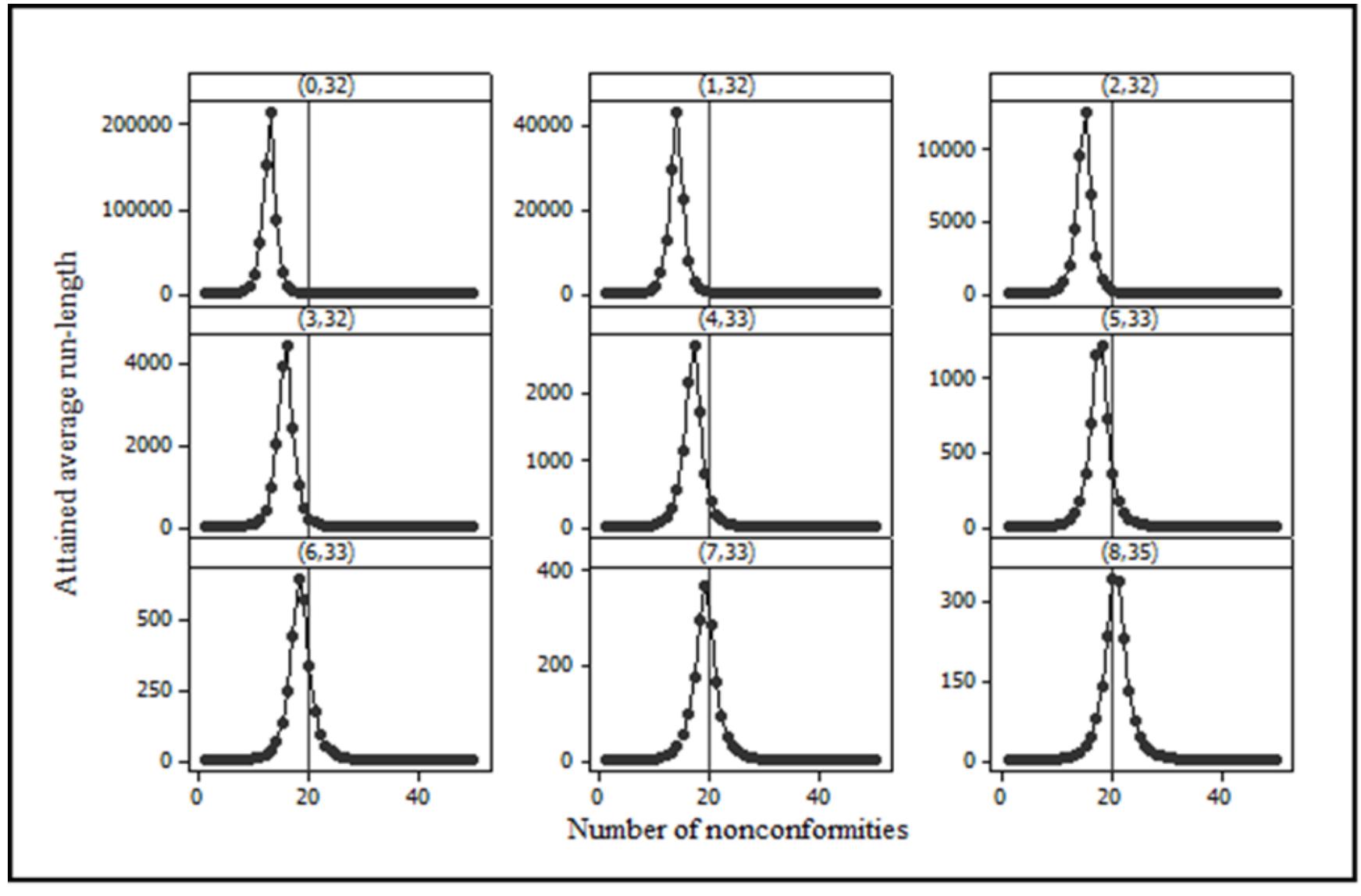
Step 3: Choose the pair $(d^\#, f^\#) \in S$ such that

$$ARL(d^\#, f^\# | c) = \min_{(d, f_i) \in S} ARL(d, f_i | c) \quad (2.27)$$

i.e. we choose the pair $(d^\#, f^\#)$ that result in the smallest value of q . Thus the nearly ARL -unbiased MIPL for the c chart are given by $LCL_c = d^\#$ and $UCL_c = f^\# + 1$. Note that, if there is more than one pair that satisfies Equation (2.27), then we must choose the pair that results in an $AARL_0$ closest to the ARL_0 .

To illustrate this, we use Example 2.2 to construct a c chart that will result in nearly ARL -unbiased control limits. Taking the pairs (d, f) in S that are given in Table 2.4 as the control limits, we construct the ARL curves shown in Figures 2.6 (a) and (b). It is evident that most of the control limits in set S have undesirable OOC performance in addition to having poor IC performance. The c chart using MIPL method with control limits $(4, 33)$ is not ARL -unbiased, however the c chart with the pair $(8, 35)$ results in $q = 0$.

(a) *ARL* curves for control limit constants in set S_1 from Table 2.3



(b) *ARL* curves for control limit constants in set S_2 from Table 3

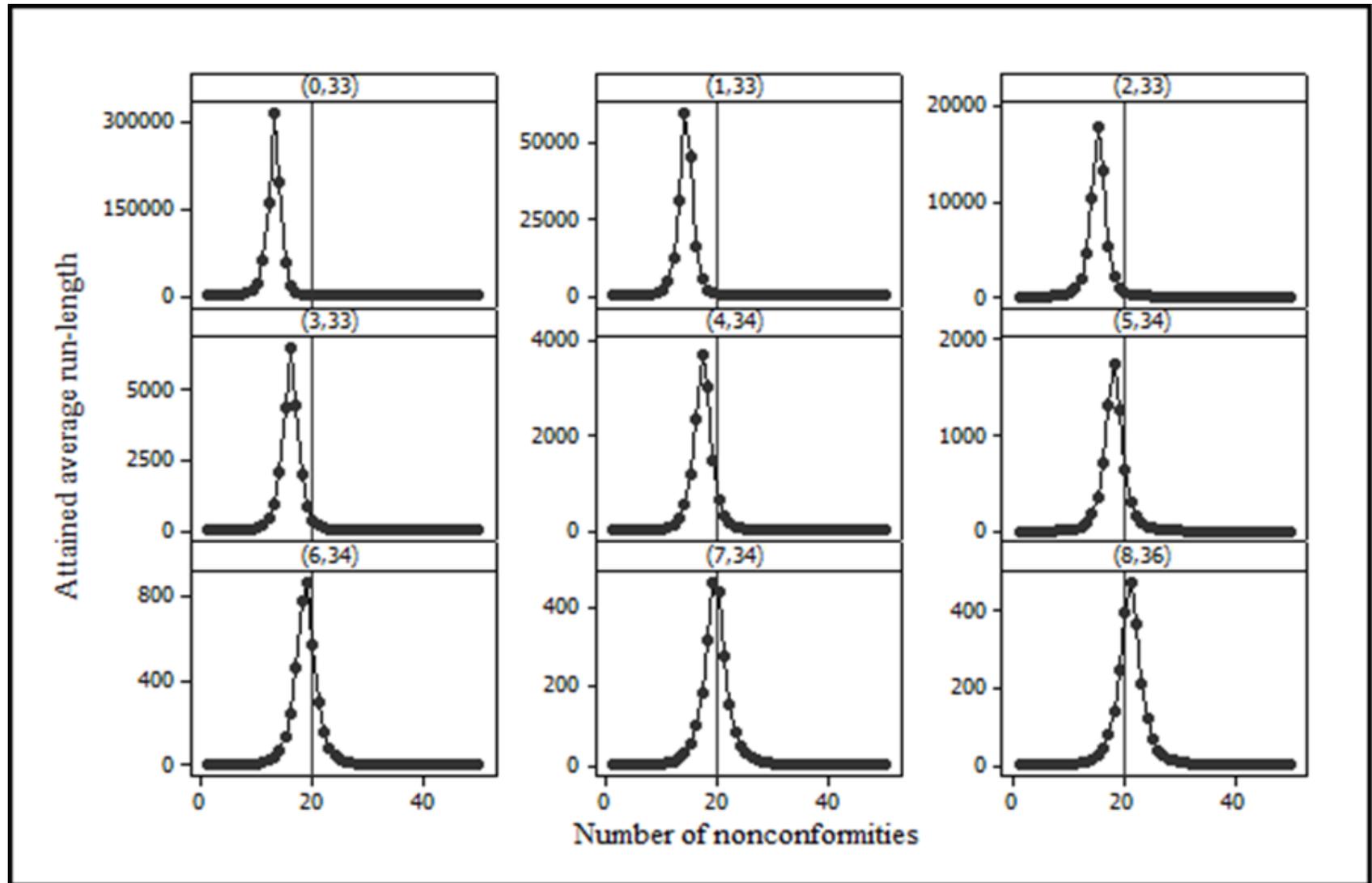


Figure 2.6: *ARL* curve for the control limit constants generated by the MIPL method for $c_0 = 20$

In Figure 2.7, we plot the competing traditional methods' *ARL* curves along with the nearly *ARL*-unbiased MIPL pair (8, 35) and deduce that the 3-SL method is *ARL*-biased. Furthermore, the CPL method has *ARL*-unbiased control limits but the $AARL_0 = 632.01$ is much higher than the expected value of 370.4 and the *ARL*-unbiased method results in *ARL*-unbiased control limits with an $AARL_0$ equal to 345.91 (much closer to 370.4 than 632.01). Therefore, if the OOC performance of the c chart is also importance, the practitioner may consider taking into account Equation (2.27) rather than Equation (2.25).

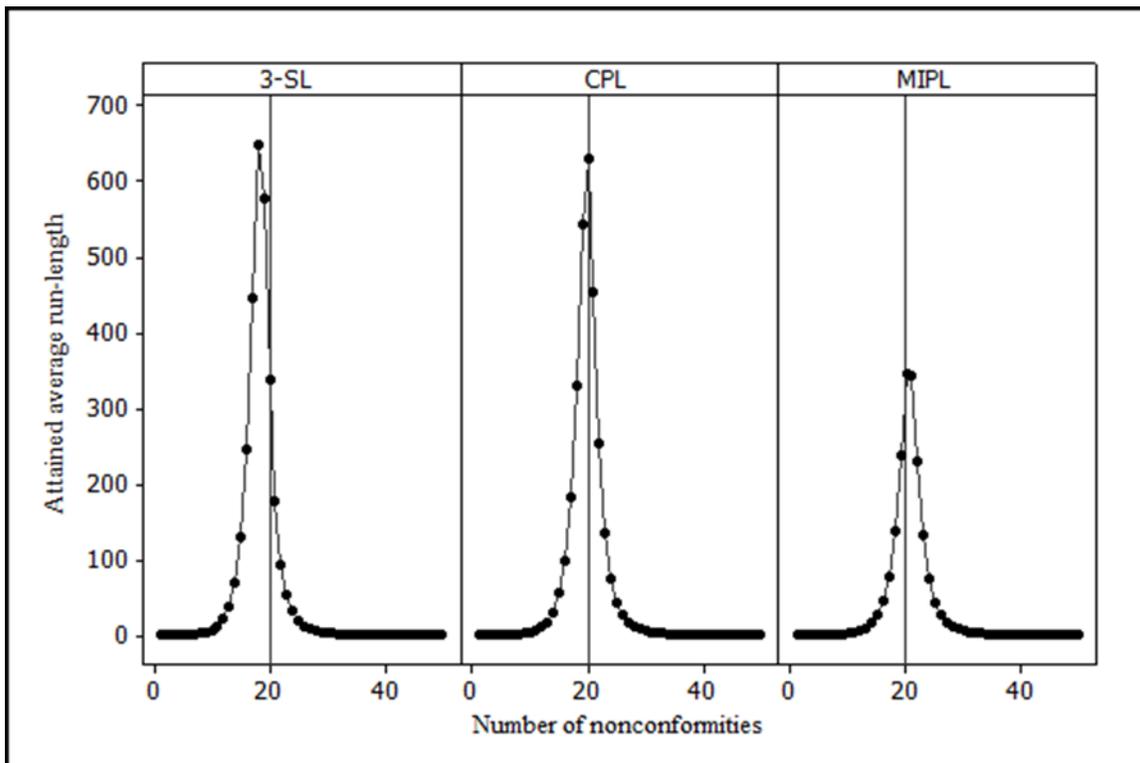


Figure 2.7: *ARL* curves for all three methods for $c_0 = 20$

Simulations indicate that when the parameter c_0 is small, the nearly *ARL*-unbiased MIPL method does not result in *ARL*-unbiased control limits, however, this method never performs worse than the 3-SL and CPL methods. In addition, we observed that as c_0 increases, the OOC MIPL method yields better performance than its competitors because the set S provides more options for the optimal pair $(d^\#, f^\#)$ compared to 3-SL and CPL methods, which have only one option for the limits (d, f) .

2.7 Concluding Remarks

In this chapter, we proposed an MIPL method to design the p and c control charts for Case K and compared the results with the traditional methods (i.e. 3-SL and CPL). The MIPL method ensures that, when the process is IC, the p and the c charts always have attained FAR and $SDRL$ values that are either the same as, or much closer to the nominal values compared to the traditional methods. Furthermore, it was shown that the MIPL approach can be formulated such that the p and c charts have similar or better nearly ARL -unbiased control limits compared to using traditional methods. The key component of the MIPL method is the fact that it creates a set of control limits that a practitioner can use to choose the best possible pair of control limits to design a p or c control chart. For very small sample sizes and/or process parameters, the MIPL method yields similar (and in some cases, better) results as the traditional methods, however as the sample size and/or process parameters increase, the MIPL method either performs similar to or outperforms the traditional methods.

2.8 Appendix 2A: Microsoft® Excel calculations

We illustrate how the results in Example 2.1 were calculated.

3-sigma limits

The 3-sigma limits calculations in Table 2.2 were calculated as follows. The formula sheet is given by,

	A	B	C	D	E	F	G	H
1				<i>a</i>	=INT(\$B\$3*E8)			
2	<i>p</i>	0.2		<i>b</i>	=IF(B3*E9=INT(B3*E9),MIN(INT(\$B\$3*E9)-1,B3),MIN(INT(\$B\$3*E9),B3))			
3	<i>n</i>	100						
4					Check if <i>b</i> is an integer			
5					=B3*E9			
6	<i>k</i>	3						
7								
8				<i>LCL</i>	=\$B\$2-B6*SQRT(\$B\$2*(1-\$B\$2)\$B\$3)			
9				<i>UCL</i>	=\$B\$2+B6*SQRT(\$B\$2*(1-\$B\$2)\$B\$3)			
10								
11								
12								
13					P(No Signal)	P(Signal)	ICARL	SDRL
14					=IF(E1<0,BINOMDIST(\$E\$2,B3,\$B\$2,TRUE),BINOMDIST(\$E\$2,B3,\$B\$2,TRUE)-BINOMDIST(\$E\$1,B3,\$B\$2,TRUE))	=1-E14	=1/F14	=SQRT(\$G\$14*(G\$14-1))
15								
16						D		
17						=ABS((F14-0.0027)/0.0027)		

and the corresponding value sheet is given by

	A	B	C	D	E	F	G	H
1				<i>a</i>	8.0			
2	<i>p</i>	0.2		<i>b</i>	31.0			
3	<i>n</i>	100						
4					Check if <i>b</i> is an integer			
5					32.00000			
6	<i>k</i>	3						
7								
8				<i>LCL</i>	0.08000			
9				<i>UCL</i>	0.32000			
10								
11								
12								
13					P(No Signal)	P(Signal)	ICARL	SDRL
14					0.99601	0.00399	250.93	250.43
15								
16						D		
17						47.60%		

Conventional probability limits

The conventional probability limits calculations in Table 2.2 were calculated as follows. The formula sheet is given by,

	A	B	C	D	E
1	n	100		upper	lower
2	p	0.2		0.00135	0.00135
3					
4	a=LCL	Left Tail		b=UCL-1	Right Tail
5	0	=BINOMDIST(A5,\$B\$1,\$B\$2,TRUE)		0	=1-BINOMDIST(D5,\$B\$1,\$B\$2,TRUE)
6	1	=BINOMDIST(A6,\$B\$1,\$B\$2,TRUE)		1	=1-BINOMDIST(D6,\$B\$1,\$B\$2,TRUE)
7	2	=BINOMDIST(A7,\$B\$1,\$B\$2,TRUE)		2	=1-BINOMDIST(D7,\$B\$1,\$B\$2,TRUE)
8	3	=BINOMDIST(A8,\$B\$1,\$B\$2,TRUE)		3	=1-BINOMDIST(D8,\$B\$1,\$B\$2,TRUE)
9	4	=BINOMDIST(A9,\$B\$1,\$B\$2,TRUE)		4	=1-BINOMDIST(D9,\$B\$1,\$B\$2,TRUE)
10	5	=BINOMDIST(A10,\$B\$1,\$B\$2,TRUE)		5	=1-BINOMDIST(D10,\$B\$1,\$B\$2,TRUE)
11	6	=BINOMDIST(A11,\$B\$1,\$B\$2,TRUE)		6	=1-BINOMDIST(D11,\$B\$1,\$B\$2,TRUE)
12	7	=BINOMDIST(A12,\$B\$1,\$B\$2,TRUE)		7	=1-BINOMDIST(D12,\$B\$1,\$B\$2,TRUE)
13	8	=BINOMDIST(A13,\$B\$1,\$B\$2,TRUE)		8	=1-BINOMDIST(D13,\$B\$1,\$B\$2,TRUE)
14	9	=BINOMDIST(A14,\$B\$1,\$B\$2,TRUE)		9	=1-BINOMDIST(D14,\$B\$1,\$B\$2,TRUE)
15	10	=BINOMDIST(A15,\$B\$1,\$B\$2,TRUE)		10	=1-BINOMDIST(D15,\$B\$1,\$B\$2,TRUE)
16	11	=BINOMDIST(A16,\$B\$1,\$B\$2,TRUE)		11	=1-BINOMDIST(D16,\$B\$1,\$B\$2,TRUE)
17	12	=BINOMDIST(A17,\$B\$1,\$B\$2,TRUE)		12	=1-BINOMDIST(D17,\$B\$1,\$B\$2,TRUE)
18	13	=BINOMDIST(A18,\$B\$1,\$B\$2,TRUE)		13	=1-BINOMDIST(D18,\$B\$1,\$B\$2,TRUE)
19	14	=BINOMDIST(A19,\$B\$1,\$B\$2,TRUE)		14	=1-BINOMDIST(D19,\$B\$1,\$B\$2,TRUE)
20	15	=BINOMDIST(A20,\$B\$1,\$B\$2,TRUE)		15	=1-BINOMDIST(D20,\$B\$1,\$B\$2,TRUE)
21	16	=BINOMDIST(A21,\$B\$1,\$B\$2,TRUE)		16	=1-BINOMDIST(D21,\$B\$1,\$B\$2,TRUE)
22	17	=BINOMDIST(A22,\$B\$1,\$B\$2,TRUE)		17	=1-BINOMDIST(D22,\$B\$1,\$B\$2,TRUE)
23	18	=BINOMDIST(A23,\$B\$1,\$B\$2,TRUE)		18	=1-BINOMDIST(D23,\$B\$1,\$B\$2,TRUE)
24	19	=BINOMDIST(A24,\$B\$1,\$B\$2,TRUE)		19	=1-BINOMDIST(D24,\$B\$1,\$B\$2,TRUE)
25	20	=BINOMDIST(A25,\$B\$1,\$B\$2,TRUE)		20	=1-BINOMDIST(D25,\$B\$1,\$B\$2,TRUE)
26	21	=BINOMDIST(A26,\$B\$1,\$B\$2,TRUE)		21	=1-BINOMDIST(D26,\$B\$1,\$B\$2,TRUE)
27	22	=BINOMDIST(A27,\$B\$1,\$B\$2,TRUE)		22	=1-BINOMDIST(D27,\$B\$1,\$B\$2,TRUE)
28	23	=BINOMDIST(A28,\$B\$1,\$B\$2,TRUE)		23	=1-BINOMDIST(D28,\$B\$1,\$B\$2,TRUE)
29	24	=BINOMDIST(A29,\$B\$1,\$B\$2,TRUE)		24	=1-BINOMDIST(D29,\$B\$1,\$B\$2,TRUE)
30	25	=BINOMDIST(A30,\$B\$1,\$B\$2,TRUE)		25	=1-BINOMDIST(D30,\$B\$1,\$B\$2,TRUE)
31	26	=BINOMDIST(A31,\$B\$1,\$B\$2,TRUE)		26	=1-BINOMDIST(D31,\$B\$1,\$B\$2,TRUE)
32	27	=BINOMDIST(A32,\$B\$1,\$B\$2,TRUE)		27	=1-BINOMDIST(D32,\$B\$1,\$B\$2,TRUE)
33	28	=BINOMDIST(A33,\$B\$1,\$B\$2,TRUE)		28	=1-BINOMDIST(D33,\$B\$1,\$B\$2,TRUE)
34	29	=BINOMDIST(A34,\$B\$1,\$B\$2,TRUE)		29	=1-BINOMDIST(D34,\$B\$1,\$B\$2,TRUE)
35	30	=BINOMDIST(A35,\$B\$1,\$B\$2,TRUE)		30	=1-BINOMDIST(D35,\$B\$1,\$B\$2,TRUE)
36	31	=BINOMDIST(A36,\$B\$1,\$B\$2,TRUE)		31	=1-BINOMDIST(D36,\$B\$1,\$B\$2,TRUE)
37	32	=BINOMDIST(A37,\$B\$1,\$B\$2,TRUE)		32	=1-BINOMDIST(D37,\$B\$1,\$B\$2,TRUE)
38	33	=BINOMDIST(A38,\$B\$1,\$B\$2,TRUE)		33	=1-BINOMDIST(D38,\$B\$1,\$B\$2,TRUE)
39	34	=BINOMDIST(A39,\$B\$1,\$B\$2,TRUE)		34	=1-BINOMDIST(D39,\$B\$1,\$B\$2,TRUE)

and the corresponding value sheet is given by

	A	B	C	D	E	F	H	I	J	K
1	n	100		upper	lower					
2	p	0.2		0.00135	0.00135					
3										
4	a=LCL	Left Tail		b=UCL-1	Right Tail		a	b		
5	0	0.00000		0	1.00000		8	33		
6	1	0.00000		1	1.00000					
7	2	0.00000		2	1.00000					
8	3	0.00000		3	1.00000					
9	4	0.00000		4	1.00000					
10	5	0.00002		5	0.99998		P(No-Signal)	P(Signal)	ICARL	SDRL
11	6	0.00008		6	0.99992		0.99841	0.00159	628.03	627.53
12	7	0.00028		7	0.99972					
13	8	0.00086		8	0.99914			D		
14	9	0.00233		9	0.99767			41.03%		
15	10	0.00570		10	0.99430					
16	11	0.01257		11	0.98743					
17	12	0.02533		12	0.97467					
18	13	0.04691		13	0.95309					
19	14	0.08044		14	0.91956					
20	15	0.12851		15	0.87149					
21	16	0.19234		16	0.80766					
22	17	0.27119		17	0.72881					
23	18	0.36209		18	0.63791					
24	19	0.46016		19	0.53984					
25	20	0.55946		20	0.44054					
26	21	0.65403		21	0.34597					
27	22	0.73893		22	0.26107					
28	23	0.81091		23	0.18909					
29	24	0.86865		24	0.13135					
30	25	0.91252		25	0.08748					
31	26	0.94417		26	0.05583					
32	27	0.96585		27	0.03415					
33	28	0.97998		28	0.02002					
34	29	0.98875		29	0.01125					
35	30	0.99394		30	0.00606					
36	31	0.99687		31	0.00313					
37	32	0.99845		32	0.00155					
38	33	0.99926		33	0.00074					
39	34	0.99966		34	0.00034					
40	35	0.99985		35	0.00015					
41	36	0.99994		36	0.00006					

Modified improved probability limits

The following formula and value sheets show how to calculate the value of L_{max} for the MIPL method in Example 2.1.

	A	B
1		
2	n	100
3	p	0.2
4		
5		
6	<i>a=LCL</i>	<i>Left Tail</i>
7	0	=BINOMDIST(A7,\$B\$2,\$B\$3,TRUE)
8	1	=BINOMDIST(A8,\$B\$2,\$B\$3,TRUE)
9	2	=BINOMDIST(A9,\$B\$2,\$B\$3,TRUE)
10	3	=BINOMDIST(A10,\$B\$2,\$B\$3,TRUE)
11	4	=BINOMDIST(A11,\$B\$2,\$B\$3,TRUE)
12	5	=BINOMDIST(A12,\$B\$2,\$B\$3,TRUE)
13	6	=BINOMDIST(A13,\$B\$2,\$B\$3,TRUE)
14	7	=BINOMDIST(A14,\$B\$2,\$B\$3,TRUE)
15	8	=BINOMDIST(A15,\$B\$2,\$B\$3,TRUE)
16	9	=BINOMDIST(A16,\$B\$2,\$B\$3,TRUE)
17	10	=BINOMDIST(A17,\$B\$2,\$B\$3,TRUE)
18	11	=BINOMDIST(A18,\$B\$2,\$B\$3,TRUE)
19	12	=BINOMDIST(A19,\$B\$2,\$B\$3,TRUE)
20	13	=BINOMDIST(A20,\$B\$2,\$B\$3,TRUE)
21	14	=BINOMDIST(A21,\$B\$2,\$B\$3,TRUE)
22	15	=BINOMDIST(A22,\$B\$2,\$B\$3,TRUE)
23	16	=BINOMDIST(A23,\$B\$2,\$B\$3,TRUE)
24	17	=BINOMDIST(A24,\$B\$2,\$B\$3,TRUE)
25	18	=BINOMDIST(A25,\$B\$2,\$B\$3,TRUE)
26	19	=BINOMDIST(A26,\$B\$2,\$B\$3,TRUE)
27	20	=BINOMDIST(A27,\$B\$2,\$B\$3,TRUE)
28	21	=BINOMDIST(A28,\$B\$2,\$B\$3,TRUE)
29	22	=BINOMDIST(A29,\$B\$2,\$B\$3,TRUE)
30	23	=BINOMDIST(A30,\$B\$2,\$B\$3,TRUE)
31	24	=BINOMDIST(A31,\$B\$2,\$B\$3,TRUE)
32	25	=BINOMDIST(A32,\$B\$2,\$B\$3,TRUE)
33	26	=BINOMDIST(A33,\$B\$2,\$B\$3,TRUE)
34	27	=BINOMDIST(A34,\$B\$2,\$B\$3,TRUE)
35	28	=BINOMDIST(A35,\$B\$2,\$B\$3,TRUE)
36	29	=BINOMDIST(A36,\$B\$2,\$B\$3,TRUE)

	A	B
1		
2	n	100
3	p	0.2
4		
5		
6	$\alpha=LCL$	Left Tail
7	0	0.0000000
8	1	0.0000000
9	2	0.0000001
10	3	0.0000006
11	4	0.0000037
12	5	0.0000187
13	6	0.0000780
14	7	0.0002770
15	8	0.0008554
16	9	0.0023336
17	10	0.0056964
18	11	0.0125749
19	12	0.0253288
20	13	0.0469122
21	14	0.0804437
22	15	0.1285055
23	16	0.1923376
24	17	0.2711890
25	18	0.3620871
26	19	0.4601614
27	20	0.5594616
28	21	0.6540332
29	22	0.7389328
30	23	0.8109128
31	24	0.8686468
32	25	0.9125246
33	26	0.9441673
34	27	0.9658484
35	28	0.9799798
36	29	0.9887510
37	30	0.9939407
38	31	0.9968703
39	32	0.9984496
40	33	0.9992631

The rest of the steps follow as discussed in Example 2.1.

Similar calculations were done for the c chart in Example 2.2.

Chapter 3

Synthetic quality control charts: An overview

3.1 Introduction

In the literature, many control charts have been proposed. Among the most popular are the Shewhart charts, the CUSUM charts and the EWMA charts. Various adaptations and generalizations of these basic charts have been considered, for example, the variable sampling interval (VSI), the variable sample size (VSS) and the double sampling (DS) charts. In this review chapter, our focus will be on a class of charts called the synthetic charts.

Wu and Spedding (2000a) originally defined a synthetic chart as the integration of a Shewhart chart and a conforming run-length (CRL) chart, see Section 1.10. Following this, Scariano and Calzada (2009) proposed a more general approach referred to as the generalized synthetic chart (GSC) procedure. For the GSC procedure, a synthetic chart is defined as the integration of some control charting procedure and a CRL chart. That is, a GSC consists of two sub-charts, one, a basic (or a classical) chart for the parameter of interest and a second, a CRL chart. However, unlike a classical chart, a signal is not based on a single charting statistic falling beyond the control limits. Instead, when any sample produces a value beyond the control limits of a classical chart, called a sub-chart of a synthetic chart, that sample is marked as nonconforming and the control procedure moves to the second sub-chart, the CRL chart, and a signal is obtained depending on the outcome of the CRL chart. Since Wu and Spedding (2000a) proposed the concept of a synthetic control chart to monitor the mean for normally distributed data, there have been a number of authors who contributed to this topic. Most, if not all, articles that discuss the concept of synthetic charts will be reviewed in this chapter. Several types of synthetic charts have been considered in the literature, with a large number of these based on the second sub-chart being the CRL chart. However, there are other synthetic-type charts where the second sub-chart is not a CRL chart but is either (i) a cumulative quantity count chart to monitor the time until the r^{th} event (denoted CQC- r); (ii) a cumulative count conforming chart to monitor the number of inspected items until the occurrence of r defects (denoted CCC- r); (iii) a group conforming run-length (these are

denoted by GCRL). In Table 3.1 we classify the synthetic chart based on the CRL sub-chart into variables (parametric and nonparametric) and attributes (parametric) control charts, respectively, for the univariate and multivariate cases, and are reviewed in Sections 3.3 and 3.4, respectively.

Table 3.1: Summary of the synthetic charts that are based on the CRL sub-chart

Data dimensionality	Synthetic chart for	Quality characteristic
Univariate	Variables: Parametric	Mean
		Variation
		Mean and Variation
		Coefficient of variation
		Mean time between events
		Economic and economic-statistical designs
	Variables: Nonparametric	Location
		Variation
	Attributes: Parametric	Fraction/number nonconforming
Average/actual number of nonconformities		
Multivariate	Variables: Parametric	Mean
		Variation
		Mean and Variation
		Economic and economic-statistical designs
	Variables: Nonparametric	Location
	Attributes: Parametric	Fraction/number nonconforming

While there are a vast number of articles based on parametric synthetic control charts for variables data, not much work has been done on parametric synthetic charts based on attributes data and nonparametric synthetic charts. In SPC there are a number of authors that have compiled literature reviews to summarize what has been done in a specific area; for example, (i) Woodall (1997) provided a broad review for parametric attributes charts; (ii) Cheng and Thaga (2006) and more recently McCracken and Chakraborti (2013) reviewed charts that jointly monitor the mean and variation; (iii) Jensen et al. (2006) and more recently Psarakis et al. (2013) investigated the effects of parameter estimation; (iv) Chakraborti et al. (2001), Chakraborti and Graham (2007) and Chakraborti et al. (2011) reviewed nonparametric charts. For synthetic charts, only Khoo (2014) did a literature review for the univariate parametric variables synthetic charts to monitor the mean and those to monitor the

variation. In this chapter, we provide a more comprehensive review of the synthetic charts following the outline in Table 3.1 as well as discussing other synthetic-type charts.

The rest of the chapter is structured as follows. In Section 3.2, the basic characteristics of the synthetic chart are discussed with an emphasis on the synthetic chart for the mean. Then in Sections 3.3 and 3.4, a review of the literature is done according to the structure outlined in Table 3.1. In Section 3.5, we briefly discuss other available synthetic-type charts in the literature, where the second sub-chart is not a CRL. Finally in Section 3.6, we give some concluding remarks with the summary of future research ideas given later in Chapter 5. In addition, some proofs of the equations that are used in this chapter are shown in Appendix 3A. In Appendix 3B, SAS® programs to calculate the run-length properties of the synthetic chart are given. A comparison study among four variables control charts is done in Appendix 3C and finally, the corresponding SAS® programs for the *2-of-2* KL and *2-of-3* KL charts are given in Appendix 3D.

3.2 Operation and run-length characteristics of the synthetic chart for the mean

As noted earlier, Wu and Spedding (2000a) proposed the first synthetic chart for the mean of variables data which is a combination of two sub-charts, the Shewhart \bar{X} chart and a CRL chart. Assuming that the observations X_{ij} 's follow a normal $N(\mu_0, \sigma_0^2)$ distribution, where μ_0 and σ_0^2 are the specified IC mean and variance, respectively, the Shewhart \bar{X} chart is the most commonly used and familiar control chart to monitor the mean. Two cases are generally considered. First, where the parameters μ_0 and σ_0^2 are known or specified, called the standards known case (i.e. Case K) and second, where the parameters are unknown/unspecified (i.e. Case U) and need to be estimated from a Phase I reference sample. Most synthetic charts are proposed for Case K but some work is available for Case U.

Moreover, Davis and Woodall (2002) showed that the run-length (RL) distribution of the synthetic chart must be obtained under two scenarios called the *zero-state* and the *steady-state* mode, respectively. In the zero-state mode, it is assumed that there is a nonconforming sample at time zero. This is known as a head-start feature and Davis and Woodall (2002) showed that it is this assumption that made the synthetic chart of Wu and Spedding (2000) seem more powerful than several popular competing charts. Zero-state is an important point,

because when it is ruled-out, as Davis and Woodall (2002) showed, the average run-length (*ARL*) performance of the synthetic chart declines. On the other hand, in the steady-state mode, one assumes that the process starts and stays IC for a long time (i.e. the effect of a head-start feature has disappeared) and then a process shift occurs at some ‘random time’. As will be discussed below, one may obtain significantly different performance results for the synthetic chart depending on what mode of analysis is assumed. We describe these in the sequel.

3.2.1 Parameters known (Case K)

The operation of the synthetic chart to monitor the mean in Case K for both the zero-state and the steady-state modes is as follows.

- Step 1. Determine the *LCL* of the *CRL* sub-chart (i.e. $H > 0$) and the distance of the control limits from the center line (i.e. $k > 0$).
- Step 2. Compute the control limits of the \bar{X} sub-chart, i.e. $UCL_{\bar{X}}/LCL_{\bar{X}} = \mu_0 \pm k \frac{\sigma_0}{\sqrt{n}}$.
- Step 3. Take a random sample of size n at each inspection time point and compute \bar{X}_i .
- Step 4. If $LCL_{\bar{X}} < \bar{X}_i < UCL_{\bar{X}}$, the sample is declared conforming and the control flow returns to Step 3. Otherwise, the control flow proceeds to Step 5.
- Step 5. Calculate the number of \bar{X} samples between the present and the last nonconforming sample. This is the plotting statistic of the *CRL* chart, denoted by CRL_i .
- Step 6. If $CRL_i > H$, the process is declared IC and the control flow returns to Step 3. Otherwise, an OOC signal is generated and the control flow proceeds to Step 7.
- Step 7. Find and remove assignable cause(s). Then return to Step 3.

The algorithm to determine the optimal values of k and H is discussed below for the zero-state and the steady-state, respectively. To illustrate the above, consider the \bar{X} sub-chart shown in Figure 1 for certain values of k , μ_0 and σ_0 . Based on the \bar{X} sub-chart, there are nonconforming samples at times, 5, 7 and 11, respectively. We determine the *CRL* plotting statistics between two nonconforming samples, first between times 0 and 5 (assuming a signal at time 0): $CRL_1 = 5$, then between times 6 and 7: $CRL_2 = 2$ and finally between times 8 and 11: $CRL_3 = 4$; see Figure 2. Now, assume that $H = 3$ (we show later how H is determined), the synthetic \bar{X} chart signals once, at time 7.

Performance properties of a control chart are typically assessed in terms of its RL distribution and the associated characteristics, such as the average (mean), median, percentiles, etc. Davis and Woodall (2002) noted that the synthetic chart is the same as a 2-of- $(H+1)$ runs-rule chart with a head-start feature. Consequently, they suggest using the Markov chain (MC) approach discussed in Champ and Woodall (1987) for the runs-rule chart which allows calculation of the entire run-length distribution along with associated characteristics such as the mean (ARL), the $SDRL$, percentiles, the cdf, etc.

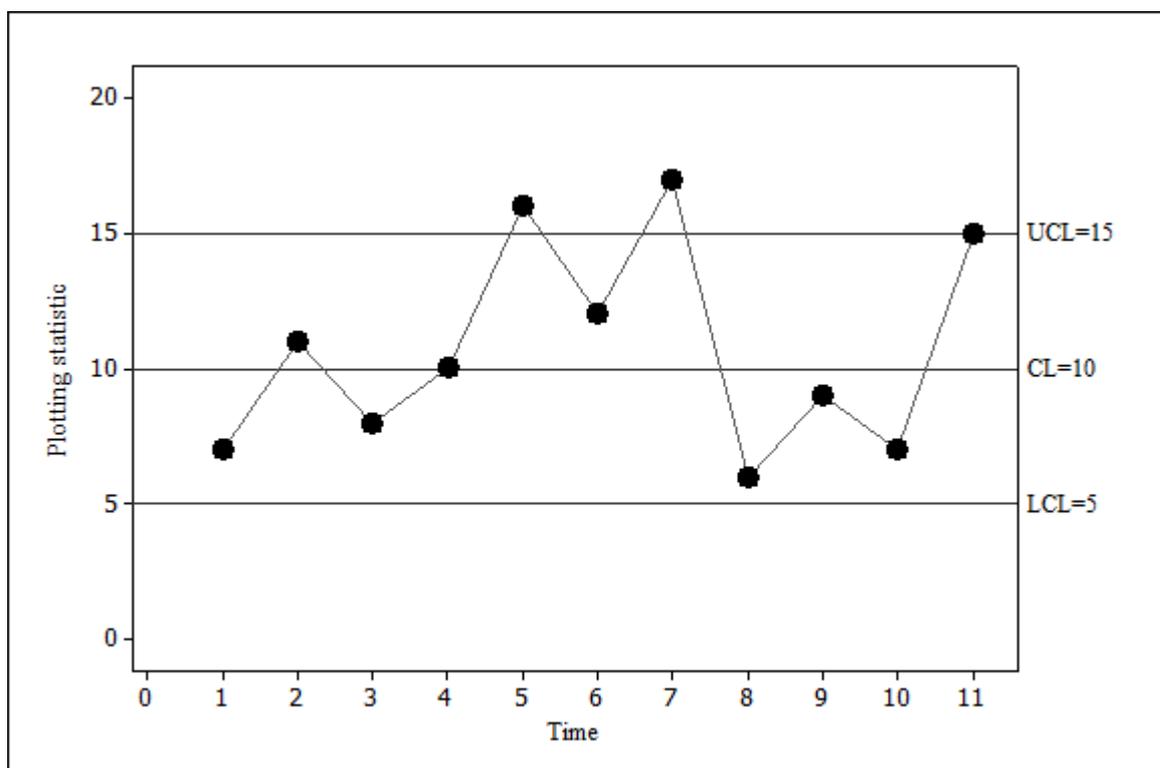


Figure 3.1: A two-sided Shewhart \bar{X} control chart

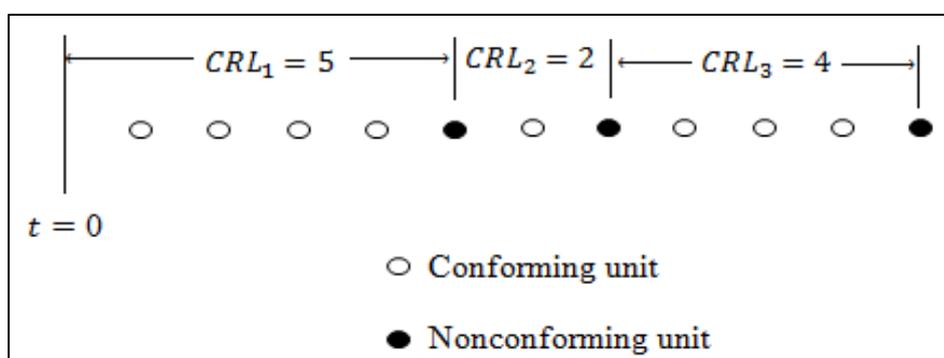


Figure 3.2: A CRL sub-chart

Davis and Woodall (2002) showed that the synthetic chart for the mean can be represented as a runs-rules chart and used an absorbing Markov chain to model the synthetic chart. Let $\theta = 1 - P(\bar{X}_i \in (LCL, UCL))$ (i.e. θ is the probability that the first sub-chart, say, the k -sigma Shewhart chart for the mean, will mark a sample as nonconforming, that is, the corresponding \bar{X}_i plots on or outside the control limits. The probability θ can be calculated exactly under the assumption of normality both in the IC and OOC cases. Then the elements of the transition probability matrix (TPM) of the Markov chain, for any general value of H are constructed as follows. For the matrix $\mathbf{Q}_{(H+1,H+1)}$, the first row contains $1 - \theta$ in the first column and θ in the second column; the last row contains $1 - \theta$ in the first column; in all other rows, the entry above the diagonal is $1 - \theta$ and in all other locations, the entry is zero. Thus, the TPM of the synthetic chart is given by

$$\mathbf{P}_{(H+2,H+2)} = \left(\begin{array}{c|c} \mathbf{Q}_{(H+1,H+1)} & \mathbf{r}_{(H+1,1)} \\ \hline \mathbf{0}'_{(1,H+1)} & \mathbf{1}_{(1,1)} \end{array} \right) \quad (3.1)$$

where $\mathbf{Q}_{(H+1,H+1)}$ is the matrix of transient probabilities given by

$$\mathbf{Q} = \begin{bmatrix} 1 - \theta & \theta & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 - \theta & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 - \theta \\ 1 - \theta & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

with $\mathbf{r}_{(H+1,1)}$ a vector that satisfies $\mathbf{r} = \mathbf{1} - \mathbf{Q}\mathbf{1}$ and a vector $\mathbf{1}_{(H+1,1)}$ is given by $(1, 1, \dots, 1)^T$, $\mathbf{0} = (0, 0, \dots, 0)^T$.

Using the seven steps “direct approach”, Wu and Spedding (2000a) showed that the *ARL* is given by (see Proof 3.4 in Appendix 3A)

$$ARL = \frac{1}{\theta(1 - (1 - \theta)^H)} \quad (3.2)$$

where $\theta = 1 - P[\bar{X}_i \in (LCL, UCL)] = 1 - \Phi(k - \delta\sqrt{n}) + \Phi(-k - \delta\sqrt{n})$. However, other RL characteristics like the cdf, pdf and *SDRL* were not provided by Wu and Spedding (2000a). We will discuss these later.

Equivalence of the synthetic chart and runs-rule chart with a head-start feature

Davis and Woodall (2002) noted that the synthetic chart for the mean is equivalent to the 2-of-($H + 1$) runs rules chart with a head-start feature (which means assuming there was a signal at time zero). To show the equivalence of the two control charts, let $H = 3$ and then follow the approach given by Davis and Woodall (2002) for a 2-of-4 chart. Suppose that each observed sample is classified as either “0” for conforming or “1” for nonconforming. Then the sequence 0101 would indicate that the second and the fourth samples are nonconforming. Let $\theta = P(\text{Next sample will be nonconforming})$, then according to Champ and Woodall (1987) the following transition probability matrix would govern the 2-of-4 control chart’s Markov chain.

	0000	0001	0010	0100		Signal
0000	$1 - \theta$	θ	0	0		0
0001	0	0	$1 - \theta$	0		θ
0010	0	0	0	$1 - \theta$		θ
0100	$1 - \theta$	0	0	0		θ
—	—	—	—	—	—	—
Signal	0	0	0	0		1

For example, assume that the Markov chain is in state “0001”, then the next sample could either be “0” or “1”. Thus, if we focus on the last four samples, the next transition could either be “0010” or “0011”. That is, the probability to go to state “0010” is $1 - \theta$, whereas the probability to go to state “0011” is θ . The rest of the Markov chain transitions can be explained similarly. Thus, taking “0001” as the initial state (assuming that the first sample plots outside the control limits i.e. with the head-start feature), then for this 2-of-4 runs-rule chart the average time for the Markov chain to eventually reach an absorbing state (i.e. gives an OOC signal) is the same as the *ARL* that results from Equation (3.2).

- **Zero-state mode**

Run-length characteristics

Using known formulas (see Fu and Lou (2003, Chapter 5)) the zero-state run-length pmf and cdf can be obtained as

$$f_{RL}(r) = \mathbf{q}^T \mathbf{Q}^{r-1} \mathbf{r} \quad (3.3)$$

$$F_{RL}(r) = 1 - \mathbf{q}^T \mathbf{Q}^r \mathbf{1} \quad (3.4)$$

respectively, for $r = 1, 2, \dots$. Note that $\mathbf{q}_{(H+1,1)}$ is the vector of initial probabilities associated with the transient states and since the second row of \mathbf{Q} corresponds to the initial state of the Markov chain then $\mathbf{q}_{(H+1,1)}$ is equal to $(0, 1, 0, \dots, 0)^T$. Furthermore, the zero-state *ARL* and *SDRL* are equal to

$$ARL = \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} \quad (3.5)$$

$$SDRL = \sqrt{2\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} - ARL^2 + ARL} \quad (3.6)$$

respectively. Using direct methods, Scariano and Calzada (2009) and Calzada and Scariano (2013b) showed that the *ARL* and *SDRL* of the synthetic chart can equivalently be written as

$$ARL = \frac{1}{\theta(1 - (1 - \theta)^H)} \quad (3.7)$$

$$SDRL = \sqrt{\frac{2 - \theta}{(1 - (1 - \theta)^H)\theta^2} + \frac{\frac{1}{\theta^2} - 2\sum_{l=1}^H l(1 - \theta)^{l-1}}{(1 - (1 - \theta)^H)^2}}. \quad (3.8)$$

respectively. Note that Expression (3.7) is the same as that first given in Wu and Spedding (2000). The advantage of (3.7) and (3.8) are that they can be calculated directly without

involving a matrix inversion. However, the advantage of the MC approach of course is that the entire run-length distribution is available.

Algorithm to determine H and k

The optimal values k and H are needed to implement the synthetic chart. In short, we determine the values of k and H that minimizes the OOC ARL , say $ARL_1(\delta_{\text{opt}})$, for a mean shift of a given magnitude expressed in units of the process standard deviation, $\delta_{\text{opt}} = |\mu_1 - \mu_0|/\sigma_0$, while ensuring that the IC ARL is equal to some given nominal value denoted by ARL_0 . Wu and Spedding (2000b) formulated the following optimization model to calculate the optimal values (and hence design the synthetic \bar{X} chart) for the zero-state:

$$\text{Minimize: } ARL_1(\delta_{\text{opt}}) = \frac{1}{\theta} \times \frac{1}{1-(1-\theta)^H}$$

$$\text{where } \theta = \theta(\delta_{\text{opt}}) = 1 - [\Phi(k - \delta_{\text{opt}}\sqrt{n}) - \Phi(-k - \delta_{\text{opt}}\sqrt{n})].$$

$$\text{Under the constraint: } \frac{1}{2\Phi(-k)} \times \frac{1}{1-[1-2\Phi(-k)]^H} = ARL_0.$$

Note that it is possible to write $\Delta = \frac{|\mu_1 - \mu_0|}{\sigma_0/\sqrt{n}}$ then $\theta = \theta(\Delta) = 1 - [\Phi(k - \Delta) - \Phi(-k - \Delta)]$ as in Scariano and Calzada (2013a) but this was not done in Wu and Spedding (2000a). We follow the method done in Wu and Spedding (2000a).

Here is how the algorithm is implemented:

- Step 1. Specify μ, σ, n, δ and ARL_0 .
- Step 2. Initialize H as 1 and find the corresponding k by solving the constraint function for the given ARL_0 .
- Step 3. Calculate $\theta(\delta_{\text{opt}})$ and $ARL_1(\delta_{\text{opt}})$ from the current k and H for the specified value of δ_{opt} .

- Step 4. For $H > 1$, compare $ARL_1(\delta_{opt})$ value for the current value of H with that of $H - 1$. If $ARL_1(\delta_{opt})$ has been reduced, increase H by one and go back to step 3 and continue iterating until increasing H no longer reduces $ARL_1(\delta_{opt})$, then go to the next step.
- Step 5. Take the k and H values corresponding to the lowest value of $ARL_1(\delta_{opt})$ as the optimal values for the synthetic chart.

Example

In Table 2, we illustrate this algorithm for $\delta_{opt} = 0.75$ and deduce that the optimal values are $k = 2.3218$ and $H = 7$ (see the boldfaced values in Table 2) for ARL_0 equal to 370.4. The Microsoft® Excel algorithm used to construct Table 3.2 is given in Appendix 3B.

Table 3.2: Values of the $ARL_1(\delta_{opt})$ for different k and H combinations when $\delta_{opt} = 0.75$ for $n = 5$ with $ARL_0 = 370.4$

H	k	$ARL_1(0.75)$
1	1.9435	6.40581
2	2.0848	5.16177
3	2.1640	4.72298
4	2.2188	4.52441
5	2.2604	4.43126
6	2.2939	4.39349
7	2.3218	4.38795
8	2.3458	4.40237
9	2.3667	4.42966
10	2.3852	4.46542

- **Steady-state mode**

Run-length characteristics

The run-length distribution of the Shewhart charts can be computed analytically, whereas the run-length distribution of other more advanced charts such as adaptive, EWMA, CUSUM and synthetic charts have to be approximated numerically since their closed form expressions cannot be obtained (see Khoo et al. (2011, 2012)). Thus, to address this problem, authors in most cases use the Markov chain. The zero-state Markov chain has already been discussed in the previous section and here we discuss the steady-state mode.

Champ (1992) simplified the procedure for computing the cyclical steady-state probability vector, denoted by \mathbf{s} , which is explained as follows:

First \mathbf{s} is computed by solving $\mathbf{b} = \mathbf{P}^T \mathbf{b}$ subject to $\mathbf{1}^T \mathbf{b} = 1$, where \mathbf{P} is the transition probability matrix with absorbing states in Equation (3.1). Then $\mathbf{s} = (\mathbf{1}^T \mathbf{z})^{-1} \mathbf{z}$ where \mathbf{z} is a vector of length $H + 1$ given by

$$\mathbf{z} = (\mathbf{G} - \mathbf{Q}^T)^{-1} \mathbf{u}$$

where $\mathbf{u} = (1 \ 0 \ 0 \ \dots \ 0)^T$ and

$$\mathbf{G} = \begin{bmatrix} 2 & 1 & 1 & \dots & 1 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Note that \mathbf{s} is called the steady-state vector that consists of the percentage of time, over many samples, that the Markov chain representing the synthetic control chart will be in each transient state, conditioned on “no signal”.

Thus, in the steady-state mode, the run-length pmf, cdf, ARL and $SDRL$ are respectively given by

$$f_{RL}(r) = \mathbf{s}^T \mathbf{Q}^{r-1} \mathbf{r} \quad (3.9)$$

$$F_{RL}(r) = \mathbf{1} - \mathbf{s}^T \mathbf{Q}^r \mathbf{1} \quad (3.10)$$

$$ARL = \mathbf{s}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} \quad (3.11)$$

$$SDRL = \sqrt{2\mathbf{s}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} - ARL^2 + ARL} \quad (3.12)$$

Algorithm to determine H and k

Similar to the zero-state mode, we determine the values of k and H that minimizes $ARL_1(\delta_{opt})$, while ensuring that the IC ARL is equal to ARL_0 . That is,

$$\text{Minimize: } ARL_1(\delta_{opt}) = \mathbf{s}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}$$

$$\text{where } \theta = \theta(\delta_{opt}) = 1 - [\Phi(k - \delta_{opt}\sqrt{n}) - \Phi(-k - \delta_{opt}\sqrt{n})].$$

$$\text{Under the constraint: } \mathbf{s}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} = ARL_0$$

$$\text{when } \theta = \theta(\delta = 0).$$

Then we follow similar steps as in the zero-state mode above.

Example

Consider Table 3, we illustrate how to obtain the optimal values of H and k for the steady-state synthetic chart for $\delta_{\text{opt}} = 0.75$. We see that the optimal values are $k = 2.2714$ and $H = 6$ (see the boldfaced values in Table 3) for ARL_0 equal to 370.4. The SAS® program to illustrate the construction of Table 3.3 is given in Appendix 3B.

Table 3.3: Values of the $ARL_1(\delta_{\text{opt}})$ for different k and H combinations when $\delta_{\text{opt}} = 0.75$ for $n = 5$ with $ARL_0 = 370.4$

H	k	$ARL_1(0.75)$
1	1.9328	8.06444
2	2.0706	7.01799
3	2.1472	6.67147
4	2.1997	6.52295
5	2.2395	6.46037
6	2.2714	6.44211
7	2.2978	6.44829
8	2.3204	6.47155
9	2.3401	6.50589
10	2.3575	6.54747

3.2.2 Parameters unknown (Case U)

For Case U, the unknown parameters μ_0 and σ_0 are estimated from an IC Phase I dataset composed of m subgroups each of n observations. Assuming independence between and within samples, a commonly used estimator for μ_0 is given by

$$\hat{\mu}_0 = \frac{1}{m} \sum_{i=1}^m \bar{X}_i$$

and a biased estimator $\hat{\sigma}_0$ for σ_0 is

$$\hat{\sigma}_0 = \sqrt{\frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2}.$$

Hence, in Case U, the control limits $L\hat{C}L$ and $U\hat{C}L$ are random variables. Khoo et al. (2008, 2009), Zhang et al. (2011) and Castagliola et al. (2013) have studied the effect of parameter estimation for the synthetic charts. Thus, the conditional probability that in Phase II, an \bar{X} sub-chart marks the i th sample as nonconforming, given the Phase I parameter estimates, is (see Zhang et al. (2011))

$$\begin{aligned} \theta_c &= 1 - P[\bar{X}_i \in (L\hat{C}L, U\hat{C}L) | \hat{\mu}_0, \hat{\sigma}_0] \\ &= \Phi\left(\frac{\sqrt{n}(\hat{\mu}_0 - \mu_0)}{\sigma_0} - k \frac{\sqrt{n}\hat{\sigma}_0}{\sigma_0} - \delta\sqrt{n}\right) + \Phi\left(-\frac{\sqrt{n}(\hat{\mu}_0 - \mu_0)}{\sigma_0} - k \frac{\sqrt{n}\hat{\sigma}_0}{\sigma_0} + \delta\sqrt{n}\right). \end{aligned} \quad (3.13)$$

Further, let $U = \frac{\sqrt{n}(\hat{\mu}_0 - \mu_0)}{\sigma_0}$ and $V = \frac{\sqrt{n}\hat{\sigma}_0}{\sigma_0}$ be the random variables with pdf $f_U(u|m) = \Phi(u|0, \frac{1}{\sqrt{m}})$ and $f_V(v|m, n) = 2vf_\gamma\left(v^2 \mid \frac{m(n-1)}{2}, \frac{2n}{m(n-1)}\right)$, respectively, where f_γ is pdf of a gamma distribution with parameters $\frac{m(n-1)}{2}$ and $\frac{2n}{m(n-1)}$. Using this estimate of θ and following Case K, one calculates first the conditional and then the unconditional RL distribution. Associated run-length distribution characteristics follow from these in a straight forward manner.

- **Zero-state mode**

Run-length characteristics

By replacing θ in Equation (3.1) by θ_c in Equation (3.13), then the conditional pmf and cdf are given by

$$f_{RL}(r) = \mathbf{q}^T \mathbf{Q}^{r-1} \mathbf{r} \quad (3.14)$$

$$F_{RL}(r) = 1 - \mathbf{q}^T \mathbf{Q}^r \mathbf{1} \quad (3.15)$$

respectively. Whereas the unconditional pdf and cdf of the RL are given by

$$f_{RL}(r) = \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{q}^T \mathbf{Q}^{r-1} \mathbf{r}) f_U(u|m) f_V(v|m, n) dv du \quad (3.16)$$

$$F_{RL}(r) = 1 - \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{q}^T \mathbf{Q}^r \mathbf{1}) f_U(u|m) f_V(v|m, n) dv du \quad (3.17)$$

respectively, see Zhang et al. (2011). Similarly, while the conditional *ARL* and *SDRL* are given by

$$ARL = \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} \quad (3.18)$$

$$SDRL = \sqrt{2\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} - ARL^2 + ARL} \quad (3.19)$$

respectively, the unconditional *ARL* (denoted by *UARL*) and the unconditional *SDRL* (denoted *USDRL*) are given by

$$UARL = \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}) f_U(u|m) f_V(v|m, n) dv du \quad (3.20)$$

$$USDRL = \int_{-\infty}^{\infty} \int_0^{\infty} \sqrt{2\mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{Q}\mathbf{1} - ARL^2 + ARL} f_U(u|m)f_V(v|m, n)dvdu \quad (3.21)$$

respectively.

Algorithm to determine H and k

For the zero-state Case U, we need to determine the values of k and H that minimizes the OOC $UARL$, say $UARL_1(\delta_{opt})$, for a mean shift of δ_{opt} , while ensuring that the IC $UARL$ is equal to some given nominal value denoted by ARL_0 . That is, the optimization model to calculate the optimal values for the zero-state for Case U:

$$\text{Minimize: } UARL_1(\delta_{opt}) = \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}) f_U(u|m)f_V(v|m, n)dvdu$$

$$\text{where } \theta_c = \theta_c(\delta_{opt}) = \Phi\left(\frac{\sqrt{n}(\hat{\mu}_0 - \mu_0)}{\sigma_0} - k\frac{\sqrt{n}\hat{\sigma}_0}{\sigma_0} - \delta\sqrt{n}\right) + \Phi\left(-\frac{\sqrt{n}(\hat{\mu}_0 - \mu_0)}{\sigma_0} - k\frac{\sqrt{n}\hat{\sigma}_0}{\sigma_0} + \delta\sqrt{n}\right).$$

$$\text{Under the constraint: } \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}) f_U(u|m)f_V(v|m, n)dvdu = ARL_0$$

$$\text{when } \theta_c = \theta_c(\delta = 0).$$

Then we follow similar steps as in the zero-state mode for Case K above. An illustration is given in Zhang et al. (2011) to determine these optimal values.

- **Steady-state mode**

Run-length characteristics

Using the steady-state vector, \mathbf{s} , it is easy to see that the conditional pmf and cdf are given by

$$f_{RL}(r) = \mathbf{s}^T \mathbf{Q}^{r-1} \mathbf{r} \quad (3.22)$$

$$F_{RL}(r) = 1 - \mathbf{s}^T \mathbf{Q}^r \mathbf{1} \quad (3.23)$$

respectively. Whereas the unconditional pdf and cdf of the RL are given by

$$f_{RL}(r) = \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{s}^T \mathbf{Q}^{r-1} \mathbf{r}) f_U(u|m) f_V(v|m, n) dv du \quad (3.24)$$

$$F_{RL}(r) = 1 - \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{s}^T \mathbf{Q}^r \mathbf{1}) f_U(u|m) f_V(v|m, n) dv du \quad (3.25)$$

respectively. Similarly, while the conditional *ARL* and *SDRL* are given by

$$ARL = \mathbf{s}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} \quad (3.26)$$

$$SDRL = \sqrt{2\mathbf{s}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} - ARL^2 + ARL} \quad (3.27)$$

respectively, the unconditional *ARL* (denoted by *UARL*) and the unconditional *SDRL* (denoted *USDRL*) are given by

$$UARL = \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{s}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}) f_U(u|m) f_V(v|m, n) dv du \quad (3.28)$$

$$USDRL = \int_{-\infty}^{\infty} \int_0^{\infty} \sqrt{2\mathbf{s}^T(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{Q}\mathbf{1} - ARL^2 + ARL} f_U(u|m)f_V(v|m, n)dvdu \quad (3.29)$$

respectively.

Algorithm to determine H and k

For the steady-state Case U, we need to determine the values of k and H that minimizes the OOC $UARL$, say $UARL_1(\delta_{opt})$, for a mean shift of a given magnitude δ_{opt} , while ensuring that the IC $UARL$ is equal to some given nominal value denoted by ARL_0 . That is, the optimization model for Case U:

$$\text{Minimize: } UARL_1(\delta_{opt}) = \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{s}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}) f_U(u|m)f_V(v|m, n)dvdu$$

$$\text{where } \theta_c = \theta_c(\delta_{opt}) = \Phi\left(\frac{\sqrt{n}(\hat{\mu}_0 - \mu_0)}{\sigma_0} - k\frac{\sqrt{n}\hat{\sigma}_0}{\sigma_0} - \delta\sqrt{n}\right) + \Phi\left(-\frac{\sqrt{n}(\hat{\mu}_0 - \mu_0)}{\sigma_0} - k\frac{\sqrt{n}\hat{\sigma}_0}{\sigma_0} + \delta\sqrt{n}\right).$$

$$\text{Under the constraint: } \int_{-\infty}^{\infty} \int_0^{\infty} (\mathbf{s}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}) f_U(u|m)f_V(v|m, n)dvdu = ARL_0$$

$$\text{when } \theta_c = \theta_c(\delta = 0).$$

Then we follow similar steps as in the zero-state mode for Case K above.

In the following Sections 3.3 and 3.4 we use the structure in Table 3.1 to review the work done for the synthetic charts for both univariate and multivariate cases.

3.3 Univariate synthetic charts

3.3.1 Parametric variables charts

A nice review for univariate synthetic charts to monitor the mean and those to monitor the variance is done in Khoo (2014). Papers already reviewed by Khoo (2014) will not be discussed but will only be mentioned briefly.

Mean

Synthetic charts to monitor the mean for skewed population were proposed in Khoo et al. (2008) and Castagliola and Khoo (2009), see the review by Khoo (2014). Moreover, Khoo (2014) reviewed the double sampling scheme proposed in Khoo et al. (2011a).

Aparisi and de Luna (2009b) formulated an optimization algorithm for no detection of shifts in a region of acceptable shifts (i.e. IC region) and at the same time, being able to detect shifts considered important (i.e. OOC region). When both the IC and OOC regions are considered, the synthetic \bar{X} chart seems to be the worst option whereas the side-sensitive synthetic \bar{X} chart (proposed in Davis and Woodall (2002)) is the best option in steady-state mode. Moreover, if the practitioner has no historical data to estimate δ , the Taguchi loss function given in Aparisi and Garcia-Diaz (2007) can be used to determine the shift region of δ to be detected and those which are not to be detected.

A synthetic chart for individual observation is an integration of the operation of the X chart and the CRL chart. Hence, Wu et al. (2010) proposed a Syn- X chart, which is a combined scheme of the synthetic X chart and the X chart. The first part of the scheme is more sensitive to small shifts, whereas the second part is more sensitive to large shifts. The Syn- X chart gives an OOC signal when either the charting statistic $y > UCL_X$ (i.e. UCL of the X chart) or $CRL < H$. Two indexes called the average extra quadratic loss (AEQL i.e. which is the weighted average ARL across the mean shift range) and the average ratio of ARL (ARARL) were used as objective functions to measure the overall performance. That is,

$$AEQL = \frac{1}{\delta_{\max}} \int_0^{\delta_{\max}} \delta^2 ARL(\delta) d\delta$$

and

$$ARARL = \frac{1}{\delta_{\max}} \int_0^{\delta_{\max}} \frac{ARL(\delta)}{ARL_{\text{Syn-X}}(\delta)} d\delta$$

respectively, where the mean shift range $0 < \delta < \delta_{\max}$, where δ_{\max} is the maximum possible mean shift. With respect to both these performance measures, it was shown that across the entire mean shift range, the Syn- X chart produces smaller OOC ARL values at more shift points or at a larger degree compared with the X chart and the synthetic X chart. Furthermore, they formulated new non-Markov chain method to derive the zero-state and steady-state ARL for the synthetic X chart and the Syn- X chart.

Zhang et al. (2011) investigated the effect of parameter estimation for the synthetic \bar{X} chart. They compared the performance of the synthetic \bar{X} chart in Case U with that in Case K. They showed that the unconditional run-length properties (i.e. ARL , $SDRL$, cdf) in Case U can be significantly different from those in Case K (especially when the number of Phase I samples is small), making it inappropriate to using the optimal parameters k and H corresponding to Case K in Case U. Moreover, they showed that in Case U, new optimal parameters (H', k') for some given m , n and δ values must be calculated so that the resulting ARL is as close as possible to the ARL of Case K. For example, when $n = 5$ and $\delta = 0.2$, the optimal values of k and H for Case K are 1.1966 and 60, respectively, so that the corresponding $ARL_1 = 127.8$ and $SDRL_1 = 167.2$. Using the same k and H in Case U scenario (with $m = 10$), these results in $UARL_1 = 801.9$ and $USDRL_1 = 880.0$, whereas the alternative parameters, $k' = 1.1449$ and $H' = 98$ results in $UARL_1 = 217.2$ and $USDRL_1 = 256.1$, for this example, these values are as close as possible to the values in Case K. In addition, tables were constructed to show how many Phase I samples are required in practice for the synthetic \bar{X} chart to have similar IC run-length distribution in both Case K and Case U.

Gan (1994) and other authors have argued that the interpretation of the run-length based on the ARL can be misleading and complicated, since the shape of the run-length distribution changes with the mean shift. Instead, the median run-length (MRL) provides a

more meaningful interpretation of the run-length for the IC and OOC performance of the charts. Hence, given the pair (H, k) , the 100γ ($0 < \gamma < 1$) percentage points of the run-length distribution corresponding to desired values of n and δ can be determined as the value m_γ such that

$$\Pr(RL \leq m_\gamma - 1) \leq \gamma \quad (3.30a)$$

$$\Pr(RL \leq m_\gamma) > \gamma \quad (3.30b)$$

If $\gamma = 0.5$ in Equations (3.30a) and (3.30b), then $m_{0.5}$ is called the median run-length (*MRL*). Thus Khoo et al. (2012) proposed a synthetic \bar{X} chart based on the percentage points of the run-length with more emphasis on the *MRL* under zero-state and steady-state modes. For both states (i.e. zero and steady) it was observed that the difference between the values of the *ARL* and the *MRL* is large when the process is IC, but it reduces as the shift increases. Furthermore, as with the *ARL*, the OOC *MRL* of the synthetic \bar{X} chart is always greater than that of the EWMA chart, unless the shift is very large.

The performance of the \bar{X} chart is based on the assumption that the mean and the standard deviation have been estimated from a homogenous retrospective samples. However, Zhang et al. (2009) showed that the \bar{X} chart suffers from wide variation from the expected IC *ARL* values, or else becomes insensitive to changes in the process mean. Hence Zhang et al. (2009) proposed a t and an EWMA- t charts that do not require estimation of the standard deviation from retrospective samples and possesses desirable robustness properties against changes in the standard deviation. Thus, Calzada and Scariano (2013a) used the GSC approach to propose the synthetic version of the t chart and the EWMA- t chart to monitor the process mean using the zero-state mode. The authors used two different methods to evaluate the synthetic charts i.e. the $ARL_1(\delta)$ based on a single specified shift value and the expected value of the OOC *ARL* (i.e. $E[ARL_1]$) over a range of shifts (see Castagliola et al. (2011)). Castagliola et al. (2011) noted that specifying a δ of interest beforehand is often too restrictive because the quality practitioner may not have historical knowledge of the process, or because shifts are not deterministic but follow some unknown distribution, i.e. it is based on $E[ARL_1]$ over the support $f_a(a)$ (i.e. the pdf of δ) rather than one specific value of δ . With respect to the zero-state *ARL* performance, the synthetic EWMA- t chart is better for

small shifts, however, for large shifts (i.e. $\delta > 1.5$), the synthetic t chart is the best. Moreover, the synthetic EWMA- t and synthetic t charts have somewhat larger $SDRL$ values for small shifts when the process is IC. However, the OOC $SDRL$ values decrease sharply for medium shifts, so that it is always less or equal to the OOC $SDRL$ values of the non-synthetic counterparts. In terms of the $E[ARL_1]$, the synthetic EWMA- t chart is optimal for small shifts whereas, synthetic t chart is optimal for large shifts. Lastly, Calzada and Scariano (2013a) noted that the synthetic EWMA- t chart is not as efficient as the synthetic EWMA- \bar{X} chart but the latter suffer from lack of robustness to estimation and changes in the process standard deviation.

Variation

A synthetic chart to monitor a sample range and a sample standard deviation based on the zero-state mode were proposed by Chen and Huang (2005) and Huang and Chen (2005), respectively. These latter two papers were reviewed in Khoo (2013). To further enhance the detection of shifts in variation, Rajmanya and Ghute (2013a, b) proposed a synthetic D chart based on the Downton's estimator under normal and non-normal data to monitor increases in the standard deviation using the zero-state mode. The Downton's estimator is given by

$$D = \frac{2\sqrt{\pi}}{n(n-1)} \sum_{j=1}^n \left(j - \frac{1}{2}(n+1) \right) X_{(j)}$$

which is an unbiased estimator of σ for normally distributed quality characteristics and $X_{(j)}$ corresponds to the order statistics of the observed data for $j = 1, 2, \dots, n$. When the underlying process distribution is normal, it was observed that the synthetic D chart produce significant ARL improvement in comparison to the R , S , D and synthetic R charts. Moreover, the performance of the synthetic D chart is similar to that of the synthetic S chart, however, for the Weibull, double exponential and gamma distributions, the synthetic D chart consistently produced smaller ARL_1 values than that of the synthetic S chart for the entire range of shifts in the standard deviation.

Joint mean and variation

In some situations, it is necessary to monitor assignable causes attributed to both the mean from μ_0 to $\mu_1 = \mu_0 \pm \delta\sigma_0$ where $\delta \neq 0$ and/or shift in the standard deviation from σ_0 to $\sigma_1 = \tau\sigma_0$ where $\tau \neq 1$, i.e. assignable causes that shift the process mean, increase the variance, or both. Hence, Costa and Rahim (2006) proposed a synthetic chart based on the non-central chi-square (NCS) statistic to jointly monitor the mean and variance using the zero-state mode. The authors stated that it is operationally simpler and more effective than the joint \bar{X} and R chart. In addition, a process is monitored using only one chart instead of looking at two charts, separately. Let $e_i = \bar{X}_i - \mu_0$ and define, $\xi_i = d$ if $e_i \geq 0$ or, $-d$ otherwise, then the NCS charting statistic is given by

$$T_i = \sum_{j=1}^n (X_{ij} - \mu_0 + \xi_i \sigma_0)^2$$

where $i = 1, 2, \dots$. When $d = 0$, $T_i/(\gamma\sigma_0)^2$ is distributed as a non-central chi-square distribution with n degrees of freedom and a non-centrality parameter $\lambda = n\delta^2/\gamma^2$. Finally, it was shown that the synthetic NCS chart with $d > 0$ is superior to the joint \bar{X} and R chart in terms of the ability to detect any kind of process disturbance.

Chen and Huang (2006) proposed a synthetic MAX chart for normally distributed data. The charting statistic is based on the $M = \max\{|U|, V\}$ where $U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}}$ and $V = \Phi^{-1} \left\{ F \left(\frac{(n-1)S^2}{\sigma_0^2} \right) \right\}$. Based on the zero-state ARL , the synthetic MAX chart was shown to be more effective than the MAX chart and the joint \bar{X} and S charts. The variable sampling interval (VSI) scheme was applied to the synthetic MAX chart and the authors showed that it is more efficient than the fixed sampling interval (FSI) scheme. Costa et al. (2008) further proposed a two-stage testing procedure, where, in the first stage, one item of the sample is inspected and if its value is close to the target value of the process mean (e.g. if $|X_{i1} - \mu_0| < w\sigma_0$, with w a design parameter of the chart) then this terminates testing, otherwise the process goes to the second stage. In the second stage, the remaining items are inspected and a NCS statistic is computed taking into account all items of the sample. For this chart the NCS statistic is given by

$$T_i = \sum_{j=1}^n (Z_{ij} + \xi_i)^2$$

where $i = 1, 2, \dots$, with $Z_{ij} = (X_{ij} - \mu_0) / \sigma_0$. Note that when $|Z_{ij}| > w$ and $T_i > k$ the i^{th} sample is classified as nonconforming. Thus, with two-stage sampling, any sample has a probability given by $\Pr[(|X_1 - \mu_0| < w\sigma_0) \cap (T_i > k)]$ to be classified as nonconforming. Moreover, the two-stage synthetic NCS chart is insensitive to decreases in the variance, except when it is followed by large shift in the mean.

Lee and Khoo (2013) proposed a synthetic mean square error (MSE) chart using the zero-state mode. Assume that T is the target value of the mean, then MSE charting statistic is given by

$$MSE_i = S_i^2 + \frac{n(\bar{x}_i - T)^2}{n-1}$$

and it follows a non-central chi-square distribution $\frac{\sigma^2}{n-1} \chi_n^2(\lambda)$ with n degrees of freedom and non-central parameter $\lambda = n \left[\frac{\mu - T}{\sigma} \right]^2$. They showed that the synthetic MSE chart performs better than the joint \bar{X} and R chart, NCS chart and synthetic NCS chart for a wide variety of shifts. Furthermore, the synthetic MSE chart always performs better than the MSE chart for all the considered shifts. In addition, for large shifts in the mean, the synthetic MSE chart consistently gives smaller ARL_1 compared to EWMA semi-circle chart (except for some cases where shift in the mean are small).

Coefficient of variation

Given that $X_{ij} \sim N(\mu_0, \sigma^2 = \omega^2 \mu_0^2)$ where $\omega = \sigma / \mu = a\omega_0$ is the population coefficient of variation (CV), a is the shift parameter and ω_0 is its IC target value. Let $W_j = \bar{X}_j / \sigma_j$ denote the charting statistic of the CV chart for Case K. Calzada and Scariano (2013b) proposed a synthetic chart to monitor increases in ω using the zero-state mode. They followed the same approach as in Calzada and Scariano (2013a) to evaluate the run-length

characteristics with respect to a single specified δ ($ARL_1(\delta)$) and using the $E[ARL_1]$ over a range of shifts. When the process is IC, the $SDRL$ of the synthetic CV chart is typically larger than that of the CV chart. However, the $SDRL_1$ and the ARL_1 of the synthetic CV chart significantly decrease for some shift δ , thereafter. With respect to both the ARL_1 and $E[ARL_1]$, the EWMA chart has better performance for small shifts of size $a < 1.5$, whereas for large shifts, the synthetic CV chart has a better performance.

Mean time-between events

Scariano and Calzada (2003) proposed a synthetic exponential chart; see the review by Khoo (2013). More recently, Yen et al. (2013) proposed three synthetic-type control charts to increase the sensitivity of the mean time between events (TBE) chart of a homogenous Poisson process. In this section, we only review the chart based on the integration of the operation of a lower one-sided Erlang's (time until the r^{th} event in a Poisson process) chart to monitor decreases in the mean TBE and the CRL chart. Using average number of observation to signal (ANOS), the authors showed that for the zero-state, the synthetic Erlang ($r = 4$) chart outperforms the exponential-EWMA chart for all shifts. Whereas for the steady-state, the synthetic Erlang ($r = 5$) chart is more efficient than the exponential-EWMA chart in detecting small to moderate shifts. For large shifts, the synthetic exponential chart and the synthetic Erlang ($r = 2$) chart perform better than the other charts.

Economic and economic-statistical designs

Yeong et al. (2012) proposed the first economic model for a synthetic chart. The authors formulated an algorithm to find the optimal parameters of the synthetic \bar{X} chart which minimizes the net sum of all costs involved, so that the chart can be operated at the economically optimal level by using the approximation of the cost function in Chung (1990). It always assumes that a process starts IC and the time until the assignable cause occurs is assumed to be exponential distributed with parameter φ . The cost function of this chart has 14 input parameters, however the only input parameters that have significant effects on the cost function are φ , δ , the quality cost per hour when the process is IC and OOC, and the cost per unit sampled. Sensitivity analysis was done by these authors, and it was further investigated in Yeong and Khoo (2013); both these papers stressed that sometimes it is not

feasible to operate the chart at the economically optimal point, hence in such a case; there are alternative parameters which can be chosen from these values that will incur minimal increase in the cost. Yeong et al. (2013a) proposed economic and economic-statistical designs under different quality loss functions and investigated the effect of misspecification of the type of the loss function, the Taguchi loss function and risk aversion coefficient of the loss function where it was shown that the penalty cost results in cases where there are larger values of δ and longer expected time to sample and interpret one unit. Also the cost function for the synthetic \bar{X} chart compares favourably with that of the \bar{X} chart and EWMA chart.

3.3.2 Nonparametric variables charts

A control chart is called nonparametric (NP) if its IC run-length distribution is the same for every continuous distribution (see e.g. Chakraborti et al. (2001)). To our knowledge, so far, only two NP synthetic control charts for Case K have been proposed in the literature, i.e. (i) based on the sign test (see Khilare and Shirke (2010)) and (ii) based on the signed-rank test (see Pawar and Shirke (2010)).

Operation and optimal values of the NP synthetic chart

To this end, let X_{ij} be some unknown continuous distribution with the i^{th} observation in the j^{th} sample. The operation of the NP synthetic chart is similar to that in Section 3.2. The optimal pair UCL^* (i.e. the UCL of the classical NP sub-chart) and H are calculated as follows. Let the IC ARL of the synthetic chart be given by $ARL(0) = \frac{1}{\theta} \times \frac{1}{1-(1-\theta)^H}$ and suppose that the desired nominal IC ARL is ARL_0 and the subgroup size is n . Then we compute the $ARL(0)$ values for all possible values of the UCL^* and $H = 1, 2, \dots$ and choose the pair UCL^* and H for which the $ARL(0)$ is closer to ARL_0 . Khilare and Shirke (2010) as well as Pawar and Shirke (2010) illustrate how to calculate these optimal values.

Location

For an ordinary sign chart, define

$$\text{sign}(X_{ij} - \mu_0) = \begin{cases} 1 & \text{if } X_{ij} > \mu_0 \\ 0 & \text{if } X_{ij} = \mu_0 \\ -1 & \text{if } X_{ij} < \mu_0 \end{cases}$$

so that the sign statistic for the i th sample is given by $SN_i = \sum_{j=1}^n \text{sign}(X_{ij} - \mu_0)$. Khilare and Shirke (2010) proposed a synthetic sign chart and they compared the zero-state ARL performance of this chart with the sign chart and the \bar{X} chart under normal, double exponential and Cauchy distributed data. They observed that the synthetic sign chart has a greater detection power for shifts in the median than the sign chart and the \bar{X} chart. In addition, they observed that the improvement in the ARL is more significant for small to moderate shifts, however, for large shifts, the ARL performance declines.

For an ordinary signed-rank (SR) chart, define

$$\psi_i = \sum_{j=1}^n \text{sign}(X_{ij} - \vartheta_0) R_{ij}^+, \text{ where } \text{sign}(X_{ij} - \vartheta_0) = \begin{cases} 1 & \text{if } X_{ij} > \vartheta_0 \\ 0 & \text{if } X_{ij} = \vartheta_0 \\ -1 & \text{if } X_{ij} < \vartheta_0 \end{cases}$$

and $R_{ij}^+ = 1 + \sum_{t=1}^n I(|x_{it} - \vartheta_0| < |x_{ij} - \vartheta_0|)$ with $I(x < y) = 1$, if $x < y$ and 0 otherwise, where ϑ_0 is target median. Pawar and Shirke (2010) proposed a synthetic SR chart and they compared the zero-state ARL performance of then chart with the $1\text{-of-}1$ SR chart, the $2\text{-of-}2$ runs-rule SR chart (see Chakraborti and Eryilmaz (2007)) and the \bar{X} chart, under three continuous and symmetric (i.e. normal, double exponential and Cauchy) distributions. They observed that the synthetic SR chart has a greater detection power for all upwards shifts in the process medians than the $1\text{-of-}1$ chart, $2\text{-of-}2$ chart and the \bar{X} chart. In addition, the zero-state ARL performance for different quartiles of the run-length distribution indicated that the synthetic SR chart had shorter interquartile ranges for all the three distributions under study.

Variation

Khilare and Shirke (2012) proposed a NP synthetic sign chart to monitor the process variation for both zero-state and steady-state modes. The charting statistics are given by $U_i = \sum_{j=1}^n U_{ij}$, with

$$U_{ij} = \begin{cases} 1 & \text{if } X_{ij} < Q_1 \text{ or } X_{ij} > Q_3 \\ 0 & \text{if } X_{ij} = Q_1 \text{ or } X_{ij} = Q_3 \\ -1 & \text{if } Q_1 < X_{ij} < Q_3 \end{cases}$$

where Q_1 and Q_3 are the 1st and 3rd quartiles, respectively. They observed that the synthetic sign chart has greater detection power for shifts in the variation than the S^2 and the sign charts for all the distributions under consideration, namely the normal, double exponential and gamma. However, the superiority of the synthetic sign chart is limited to the zero-state, since in the steady-state, the *ARL* performance declines. Furthermore, the authors proposed a NP side-sensitive synthetic control chart to monitor variation based on quartiles and they observed that side-sensitive synthetic chart performs better than the synthetic sign chart for all distributions that were considered.

3.3.3 Parametric attributes charts

Attributes control charts are needed in situations when data consist of qualitative information, as variables control chart cannot be used. For illustration of the operation of the attributes synthetic chart, see Chapter 4.

Fraction/number nonconforming and average/actual number of nonconformities

Hence Wu et al. (2001) presented the design, operation and zero-state average time to signal (*ATS*) performance for the synthetic np chart. In addition, Wu and Yeo (2001) provided the description of the algorithm that can be used to obtain optimal chart parameters UCL_{np} and H of the synthetic np chart that minimizes the OOC *ATS* for some given shift, denoted by $ATS(p_\delta)$, for the input parameters p_0 (IC fraction nonconforming), p_δ (OOC fraction nonconforming) and τ (i.e. the actual IC *ATS*), while ensuring that $ATS(p_0) \geq \tau$.

Furthermore, Wu et al. (2001) showed that the synthetic np chart performs better than the CRL chart and the np chart, also for some shifts, the $ATS(p_\delta)$ of the synthetic chart is 50% lower than that of the np chart. However, Bourke (2008) showed that the apparent superior performance reported in Wu et al. (2001) is due to a limited choice of circumstances for making comparisons. Bourke (2008) showed that in the zero-state mode, the geometric CUSUM chart (see Bourke (1991)) outperforms the synthetic np chart. In addition, the RL_2 (moving sums of successive pairs of CRLs; see Bourke (1991)) is somewhat inferior to the synthetic np chart when the $ATS(p_0)$ is set at a relatively high value of 10000, however, when the $ATS(p_0)$ is lowered to 5000, the RL_2 performs better than the synthetic np chart. In the steady-state mode, the advantage of the synthetic np chart over the np chart is at most 3%, which is not significant enough to adopt the more complicated synthetic chart. Finally, the geometric CUSUM and the RL_2 charts always perform better than the synthetic np chart in steady-state mode.

Following Wu et al. (2010), Haridy et al. (2012) proposed a combined scheme of the synthetic np chart and the np chart. In addition, they adopted the non-Markov chain approach to calculating the steady-state $ATS(p_\delta)$ of the chart. The authors used the index $WAATS$ (weighted average $ATS(p_\delta)$ produced across a range of fraction nonconforming, p_δ , $p_0 < p_\delta \leq p_{max}$) given by

$$WAATS = \frac{1}{\delta_{max} - 1} \sum_{\delta=2}^{\delta_{max}} p_\delta \times ATS(p_\delta)$$

to measure the overall performance of the combined scheme chart and they showed that the combined scheme is more effective than the np chart and the synthetic np chart by 73% and 13%, respectively, in terms of the $WAATS$.

Castagliola et al. (2013) studied the effect of parameter estimation for synthetic attributes charts. An algorithm in Section 3.2 was used to obtain the optimal parameters k and H for shifts of sizes $\delta = 0.25, 0.75$ and 1.5 for small, medium and large shifts, respectively. Then, using these parameters and the normal approximation to the binomial and Poisson distributions, respectively, they evaluated the run-length properties of the synthetic p , np , c and u charts using the Markov chain approach for both Case K and Case U. Note that this is

an attributes charts version of what was done in Zhang et al. (2011) for the synthetic \bar{X} chart. Firstly, the authors showed that the IC *ARL* and the IC *SDRL* can be significantly different to the corresponding nominal values for both Case U and Case K, more especially when the number of Phase I samples (m) is small in Case U. Thus, it seemed inappropriate to use the optimal values of k and H for Case K in the Case U scenario. Hence they suggested alternative chart parameters $k' (\neq k)$ and $H' (\neq H)$ which takes m , n and δ as given, then computes the alternative parameters such that the IC *ARL* value corresponding to Case U is as close as possible to the IC *ARL* value for Case K. In addition, they gave an indication of how many Phase I samples are required for a chart in Case U to have a similar IC performance to that of Case K. More recently, Chong et al. (2014) proposed a DS scheme for the synthetic np chart.

3.4 Multivariate synthetic charts

In many applications the data are multivariate and need to be monitored on p variables. Let the data vectors \mathbf{X}_j , where $j = 1, 2, \dots, n$ be a random sample from a p -variate distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$. These vectors represent measurements on p process characteristics or variables. Multivariate control charts have been developed in the literature to enable joint monitoring of more than one quality characteristic simultaneously, instead of using separate classical charts for each of the quality characteristics. See Bersimis et al. (2007) for a review of multivariate control charts. Synthetic version of some of these charts have been developed. Note that the operation and the algorithm to obtain the optimal design parameters k and H are similar as those for the synthetic \bar{X} chart, however, in this case δ is called the Mahalanobis distance used to measure a change in the process mean vector. Assume that $\bar{\mathbf{X}}$ denotes a $(p \times 1)$ vector of sample means, $\boldsymbol{\Sigma}_0$ denotes a $(p \times p)$ known IC covariance matrix and $\boldsymbol{\mu}_0$ denotes a $(p \times 1)$ known IC mean vector. A multivariate version of the synthetic \bar{X} chart was first proposed by Ghute and Shirke (2008b) and it is an integration of the operation of (Hotelling's) T^2 sub-chart and the CRL sub-chart. The charting statistic of synthetic T^2 chart is given by

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

The sample statistics are plotted against an upper control limit, $UCL = \chi_{\alpha}^2(p)$. That is, when the process is IC, T^2 has a central chi-square distribution with p degrees of freedom. However, when the process is thought to be OOC then T^2 has a non-central chi-square distribution with a non-central parameter equal to $\lambda^2 = n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0) = n\delta^2$.

3.4.1 Parametric variables charts

Mean

Ghute and Shirke (2008b) observed that the synthetic T^2 chart consistently outperforms the T^2 chart and the T^2 with runs-rules chart for the entire range of shifts in the process mean vector. Also, for large samples sizes and large shifts, the performance of the synthetic T^2 chart and the Hotelling's T^2 chart are similar. Note, Ghute and Shirke (2008b)

only evaluated the *ARL* for zero-state mode and only considered large shifts. Aparisi and de Luna (2009a) considered both the zero-state and steady-state cases as well as small shifts. Similar to univariate synthetic charts, the zero-state *ARL* performance is shown to be always better than that of the steady-state. Furthermore, for very small shifts, the synthetic T^2 chart cannot compete with the multivariate EWMA (MEWMA) chart, the VSS- T^2 chart and the DS- T^2 chart.

In Aparisi and de Luna (2007), a special synthetic T^2 control chart was proposed, such that it does not detect small shifts in the acceptable region (i.e. IC region) but ensures a good performance in detecting moderate and large shifts (i.e. OOC region). The chart entails specifying two objective functions i.e. first to maximize the *ARL* in the IC region and the second to minimize the *ARL* in the OOC region. The authors demonstrated that it is possible to formulate a zero-state multi-objective optimization using a generic algorithm to find the Pareto-optimal front of non-dominated solutions for the optimization problem, so that only moderate and large shifts can be detected quickly. Following the univariate synthetic weighted standard deviation (WSD) method coined by Khoo et al. (2008), same authors proposed a synthetic T^2 chart based on the WSD method aimed to improve the sensitivity of the T^2 chart in monitoring (known and unknown) multivariate skewed distributions in 2009, using the zero-state mode. Note that when the underlying process distribution is symmetric, the synthetic WSD- T^2 chart reduces to a synthetic T^2 chart proposed in Ghute and Shirke (2008b). For both Case K and Case U, the authors noted that when the process is IC, the synthetic WSD- T^2 chart gives lower *FARs* compared to the WSD- T^2 chart, the T^2 chart and the WSD-EWMA chart when the underlying process is skewed. Furthermore, for OOC cases, it has the highest mean shift detection rates among all the charts for skewed populations, when moderate and large shifts are of interest based on various values of the skewness. For multivariate charts, Khoo et al. (2011b) proposed a synthetic T^2 chart based on the percentage points of the run-length with more emphasis on the *MRL* (this is the multivariate version of work done in Khoo et al. (2012)). It was shown that the zero-state *MRL* performance surpasses that of the steady-state. Furthermore, the zero-state mode synthetic T^2 chart outperforms the MEWMA chart when $\delta \geq 1$. For steady-state, the OOC *MRL* values of the synthetic T^2 chart are greater than that of the MEWMA chart, unless the shift is large. For small shifts (in both states), the T^2 chart was observed to have the worst performance.

Lee (2012) used the GSC procedure to propose a synthetic MEWMA chart using the zero-state mode. The charting statistic is given by

$$T^2 = \mathbf{Z}_j' \boldsymbol{\Sigma}_{Z_j}^{-1} \mathbf{Z}_j$$

where $\boldsymbol{\Sigma}_{Z_j} = \left(\frac{r}{2-r}\right) \boldsymbol{\Sigma}_X$, with $\boldsymbol{\Sigma}_X$ an identity matrix and $\mathbf{Z}_j = r\mathbf{X}_j + (1-r)\mathbf{X}_{j-1}$ with \mathbf{Z}_0 a zero vector and \mathbf{X}_j is a vector of p quality characteristics observed at j^{th} sample. The authors show that the synthetic MEWMA chart is always faster than the MEWMA chart and the synthetic T^2 chart in detecting shifts, as well as the T^2 chart (when the sample size is small). However, for very large n , the synthetic MEWMA chart has a better performance than the T^2 chart for small and moderate shifts. Lee et al. (2013) further used the GSC procedure to propose a synthetic multivariate CUSUM (MCUSUM) chart for detecting shifts in the mean vector using the zero-state mode. Let \mathbf{Z}_t be the standardized sample mean, that is, independently normal, each with p different quality components of interest. The charting statistic is given by

$$T_t = [\mathbf{S}_t' \boldsymbol{\Sigma}_Z^{-1} \mathbf{S}_t]^{1/2}$$

with

$$\mathbf{S}_t = \begin{cases} (\mathbf{S}_{t-1} + \mathbf{Z}_t - \boldsymbol{\mu}_{0Z}) \left(1 - \frac{k}{C_t}\right) & \text{if } C_t > k \\ 0 & \text{otherwise} \end{cases}$$

where $C_t = \{(\mathbf{S}_{t-1} + \mathbf{Z}_t - \boldsymbol{\mu}_{0Z})' \boldsymbol{\Sigma}_Z^{-1} (\mathbf{S}_{t-1} + \mathbf{Z}_t - \boldsymbol{\mu}_{0Z})\}^{1/2}$ where $k > 0$ is the reference value and $\boldsymbol{\Sigma}_Z$ is the covariance matrix in correlation form. The authors noted that, in practice, it is impossible to specify the exact size of the shift in the mean vector, hence; they proposed an index to evaluate the overall performance of the chart over a range of pre-specified shifts i.e. the average ratio of steady-state ARL ($ASARL$) which is given by

$$ASARL = \frac{1}{h+1} \sum_{i=0}^h \frac{SARL \text{ of compared chart at } \delta_i}{SARL \text{ of MCUSUM}_{\text{Synth}} \text{ chart at } \delta_i}$$

where h is the number of steady-state ARL ($SARL$) for the control chart and $\delta_i = \delta_1 + \frac{i}{h}(\delta_2 - \delta_1)$, where $[\delta_1, \delta_2]$ is the pre-specified shift range. Thus, with respect to the $ASARL$,

the synthetic MCUSUM chart performs better than the MCUSUM, the synthetic T^2 and the T^2 charts in detecting a specified range of shifts. Also, it performs slightly better than the synthetic MEWMA chart in detecting moderate to large shifts.

The methods reviewed so far for the multivariate synthetic charts were for fixed sampling scheme (FSS) scheme. The first VSS scheme was proposed by Khoo et al. (2013) namely a synthetic T^2 chart based on the DS method using the zero-state mode. The synthetic DS chart operation is more complicated since a sample is classified as conforming or nonconforming depending on the information given by not only the first sample, but the joint information on the first and second sample. In addition, the DS synthetic chart requires calculation of six parameters for some given δ , whereas the FSS synthetic chart requires only two (H and k). Furthermore, the authors observed that the synthetic DS- T^2 chart performs better than the DS- T^2 chart for almost all shifts based on the ARL_1 and the $ANOS_1$. Moreover, the MEWMA chart only outperforms the synthetic DS- T^2 chart for small shifts of size $\delta \leq 0.5$.

Variation

Ghute and Shirke (2008a) proposed a synthetic $|S|$ chart to monitor shifts in the covariance matrix of bivariate and multivariate processes. This chart is based on the use of the determinant of the sample covariance matrix called the generalised variance, i.e.

$$|S| = \left| \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' \right|.$$

The (zero-state) ARL performance is based on the determinant ratio (DR) of the IC and OOC covariance matrix i.e. $DR = |\Sigma_0|/|\Sigma_1|$, where a value equal to one implies the process is IC, however, when $DR > 1$ or $DR < 1$, the dispersion variable is thought to have increased or decreased, respectively. For the bivariate case, the authors observed that the synthetic $|S|$ chart outperforms the $|S|$ chart, the adaptive sample size $|S|$ chart and the bivariate EWMA chart. For the multivariate case, the synthetic chart consistently produces smaller ARL_1 values than the multivariate $|S|$ chart. Note that Lee and Khoo (2013) applied the VSI scheme to the synthetic $|S|$ chart. In Machado et al. (2009a) a synthetic VMAX chart, to monitor the

covariance matrix of two correlated quality characteristics, say X and Y , is proposed, assuming that the underlying process is a bivariate normal distribution with the analysis done for both zero-state and steady-state. The points plotted on this chart correspond to the maximum of the values of the two variances i.e. $VMAX = \max(S_x^2, S_y^2)$. The IC and OOC covariance matrix are given by

$$\Sigma_0 = \begin{pmatrix} \sigma_x^2 & \sigma_{yx} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \text{ and } \Sigma_1 = \begin{pmatrix} a.a.\sigma_x^2 & a.b.\sigma_{yx} \\ a.b.\sigma_{xy} & b.b.\sigma_y^2 \end{pmatrix},$$

respectively. They observed that the synthetic VMAX chart is more efficient than the $|S|$ chart. In addition, it has a similar performance to that of the VMAX chart, except for situations where H is small and the assignable cause(s) affects only one variable (i.e. $a = c$ & $b = 1$ or $a = 1$ & $b = c$).

Joint mean and variation

Ghute and Shirke (2007) proposed the use of the combined scheme involving simultaneous use of the synthetic T^2 and the synthetic $|S|$ chart for normally distributed data. It was found that, as the mean shifts and/or the variability in the covariance matrix increases, the zero-state ARL performance comparison indicated that the combined synthetic scheme performs better than the combined T^2 and $|S|$ chart for the entire range of shifts in the process parameters. Machado et al. (2009b) used the concept of VMAX, introduced in Machado et al. (2009a), to propose a synthetic MVMAX chart to jointly monitor the mean vector and covariance matrix of a bivariate process. Consider the two correlated variables (X, Y) with sample means (\bar{X}, \bar{Y}) and sample variances (S_x^2, S_y^2) . Then the charting statistic of this chart is given by $V = \max\{|Z_x|, |Z_y|, W_x, W_y\}$, where $Z_x = \sqrt{n}(\bar{X} - \mu_x)/\sigma_x$, $Z_y = \sqrt{n}(\bar{Y} - \mu_y)/\sigma_y$, $W_x = kS_x/\sigma_x$ and $W_y = kS_y/\sigma_y$. The parameter k is required to attend the imposed condition that during the IC period, the four statistics have the same probability to exceed the UCL of the MVMAX sub-chart. The authors showed that the synthetic MVMAX chart is faster than the NCS chart and the joint T^2 and $|S|$ chart except when the correlation between X and Y is high. Moreover, the authors noted that the higher the correlation the better the performance of joint T^2 and $|S|$ chart.

Economic and economic-statistical designs

Yeong et al. (2014a) proposed an economic design and an economic-statistical design for the synthetic T^2 chart. An algorithm to find the optimal parameters of the synthetic T^2 chart, which minimizes the net sum of all the costs involved, was derived so that the chart can be operated at an economically optimal level. For this chart the cost function consists of 15 input parameters and the p quality characteristics follow a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The assignable cause is assumed to be exponential distributed with parameter φ . The authors observed that the synthetic T^2 chart yields lower minimum cost than the T^2 chart for all examples, under both economic and economic-statistical designs. Furthermore, the economically optimal design for the T^2 chart results in weaker statistical performance compared to the synthetic T^2 chart. Furthermore, they showed that the optimal parameters are quite robust to changes in the input parameters, except p and δ , thus care should be taken when approximating these. Moreover, the synthetic T^2 chart outperforms the MEWMA chart, except for small δ and when a long time is required to sample and interpret one unit. The multivariate version of the work discussed in Yeong et al. (2013) and Yeong and Khoo (2013) has been done in Yeong et al. (2014b) and Yeong et al. (2014c), respectively.

3.4.2 Nonparametric variables charts

Location

Bennett (1964) proposed a NP bivariate SR statistic which is computed for each variate in $X_j = (X_{1j}, X_{2j})$ using n observations in a sample. Let T_1 and T_2 be the two signed-rank statistics corresponding to two variables. For the i^{th} variate, define the signed-rank statistic

$$T_i = \sum_{j=1}^n C(X_{ij}) R(X_{ij}) \quad \text{for } i = 1, 2$$

where

$$C(X_{ij}) = \begin{cases} 1 & \text{if } X_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$$

and $R(X_{ij})$ is the rank of $|X_{ij}|$ among $|X_{1i}|, |X_{2i}|, \dots, |X_{ni}|$. Let $E(T_i|\boldsymbol{\mu} = \boldsymbol{\mu}_0) = v_i$ for $i = 1, 2$. Then, the Bennett's SR statistic is given by

$$W = (\mathbf{T} - \mathbf{v})' \hat{\boldsymbol{\beta}}^{-1} (\mathbf{T} - \mathbf{v})$$

where, $\mathbf{T} = (T_1, T_2)'$, $\mathbf{v} = (v_1, v_2)'$ and $\boldsymbol{\beta} = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}$ is the variance-covariance matrix of the vector \mathbf{T} . When $\boldsymbol{\mu} = \boldsymbol{\mu}_0$, the statistic W is asymptotically $\chi^2(2)$ distributed. The chart indicates that a process has shifted when $W > UCL^*$. Thus, Ghute and Shirke (2012) developed a control chart based on the integration of the operation of the bivariate SR chart and a CRL chart. They observed that the $ARL_1(\delta)$ of the bivariate synthetic SR chart is smaller than that of the T^2 chart for small shifts in the location, whereas for large shifts the performance of the bivariate synthetic SR chart is equivalent to that of the T^2 chart. Under a heavy tail bivariate double exponential distribution, the bivariate synthetic SR chart is uniformly better than the T^2 chart. However, when the data is normal, then the two charts have an equivalent performance.

3.4.3 Parametric attributes charts

Fraction/number nonconforming

In multi-attribute process, a product has different types of defects. A single multi-attribute chart can be used to monitor the number of nonconforming units found in a sample of size n with respect to the q attributes (or defect types). In this case, a unit is classified as nonconforming if any of the q types of defects is found. Haridy et al. (2013) used the multi-attribute np chart design discussed in Jolayemi (1999) to propose a multi-attribute synthetic np chart and a combined scheme of the multi-attribute synthetic np chart and multi-attribute np chart (i.e. a multi-attribute version of the scheme in Haridy et al. (2012)). They used the index AND (average number of defectives produced across a range of fraction nonconforming, $p_\delta, p_0 < p_\delta \leq p_{max}$) given by

$$AND = \frac{1}{\delta_{max} - 1} \sum_{\delta=2}^{\delta_{max}} p_{\delta} \times ATS(p_{\delta})$$

to measure the overall performance of the control chart. They observed that the combined scheme is more effective than multi-attributes np chart and synthetic np chart by 83% and 27%, respectively, in terms of AND .

3.5 Other synthetic-type charts

All the synthetic charts reviewed to this point were based on the second sub-chart being the CRL. However, other synthetic-type charts with a different second sub-chart have been proposed in the literature. To our knowledge, three such other synthetic-type charts exist and these have a second sub-chart called the GCRL, CQC- r and CCC- r based on the geometric, gamma and negative binomial distributions, respectively. Consider the GCRL sub-chart; this chart is similar to the CRL sub-chart, except in the decision making procedure. The GCRL gives an OOC signal when the first CRL charting statistic is less or equal to H or any two consecutive CRL charting statistics are both less than or equal to H for the first time. That is, when $CRL_1 \leq H$ or $CRL_i \leq H$ & $CRL_{i+1} \leq H$ for $i = 2, 3, \dots$, then the process is considered OOC. Suppose that defects in a process occur according to a Poisson distribution with parameter c per unit quantity of product. Then the number of units required to observe exactly r (≥ 2) defects has an Erlang or gamma distribution. Hence the chart to monitor TBE until the r^{th} event in a Poisson process is called the CQC- r (see Fang et al. (2013)). The CCC- r chart is based on the quality characteristic to monitor the cumulative count of items inspected until observing r (≥ 2) nonconforming items (see Mishima et al. (2002)).

Fang et al. (2013) proposed a synthetic exponential chart and a synthetic Erlang chart by integrating the operation of an exponential chart and an Erlang chart, respectively, with a GCRL chart. These synthetic charts were shown to perform better than their non-synthetic counterparts in detecting mean shifts of all sizes for both zero-state and steady-state modes. Kusakawa and Ohta (2005) proposed a synthetic confirmation sample (CS) chart for a high yield process by integrating the operation of the CS (proposed in Steiner (1999)) and the CQC- r chart. Two additional synthetic charts for high yield processes with a CCC- r sub-chart were proposed in Mishima et al. (2002) and in Kusakawa et al. (2008) with a CS sub-chart and an EWMA sub-chart, respectively. All three of these latter synthetic charts were shown to have better performance than their non-synthetic counterparts.

3.6 Concluding remarks

In this chapter, the synthetic charts for variables (parametric and nonparametric) and attributes (parametric) for univariate and multivariate data were reviewed. It is worth noting that a significantly large part of this review is based on variables synthetic charts whereas synthetic charts based on discrete and nonparametric distribution is little. For synthetic charts, it is important for a practitioner to indicate (clearly) whether zero-state or steady-state analysis was used to evaluate the run-length distribution since significantly different performance results are obtained depending on what state is assumed. For most comparisons done using the zero-state, the synthetic chart outperformed the non-synthetic counterparts for various shifts. This is because of the implicit head-start feature given to this method, which gave it a large overall zero-state advantage. When ruled out, the run-length performance deteriorates. Therefore, care needs to be taken when making conclusions about the performance of a synthetic control chart.

3.7 Appendices

3.7.1 Appendix 3A: Proofs of Equations

The proofs of the results and equations given in Chapter 3 are provided in this appendix.

Proof 3.1 - Equation (3.1)

Note that the proof for Equation (3.1) follows directly from the example illustrating the equivalence of the runs-rules chart and the synthetic chart; see Davis and Woodall (2002).

Proof 3.2 - Equation (3.3 & 3.4)

Kritzinger (2011, p. 213) showed, using the methodology in Fu et al. (2002), that

$$P(RL = r) = \boldsymbol{\xi} \mathbf{Q}^{r-1} (\mathbf{I} - \mathbf{Q}) \mathbf{1} \quad (\text{A3.1})$$

where $\boldsymbol{\xi}$ is the initial vector associated with TPM. Thus taking $\boldsymbol{\xi} = \mathbf{q}^T$ yields

$$P(RL = r) = f_{RL}(r) = \boldsymbol{\xi} \mathbf{Q}^{r-1} (\mathbf{I} - \mathbf{Q}) \mathbf{1} = \mathbf{q}^T \mathbf{Q}^{r-1} \mathbf{r} \quad (\text{A3.2})$$

as given in Equation (3.3). Moreover, Kritzinger (2011, p. 220) showed that

$$P(RL \leq r) = 1 - \boldsymbol{\xi} \mathbf{Q}^r \mathbf{1}. \quad (\text{A3.3})$$

Similarly, it follows that Equation (3.4) is given by

$$P(RL \leq r) = F_{RL}(r) = 1 - \mathbf{q}^T \mathbf{Q}^r \mathbf{1}. \quad (\text{A3.4})$$

Proof 3.3 - Equation (3.5 & 3.6)

Fu and Lou (2003, p. 73) proved that the first moment is given by $\boldsymbol{\xi}(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}$. Thus taking $\boldsymbol{\xi} = \mathbf{q}^T$ yields

$$ARL = \boldsymbol{\xi}(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} \quad (\text{A3.5})$$

as given in Equation (3.5). Furthermore, Kritzinger (2011, pp. 215-216) derived the second moment using the method in Fu and Lou (2003, Chapter 5) and showed that the standard deviation of the run-length is equal to

$$SDRL = \sqrt{\mathbf{q}^T(\mathbf{I} + \mathbf{Q})(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{1} - (\mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1})^2}. \quad (\text{A3.6})$$

However, in Latouche and Ramaswani (1999) Equation (A3.9) is given by

$$SDRL = \sqrt{2\mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{Q}\mathbf{1} - (ARL)^2 + ARL}. \quad (\text{A3.7})$$

Hence, our aim here is to show that Equations (A3.6) and (A3.7) are the same. Accordingly, we use the moment generating function in Fu et al. (2002):

$$M(t) = (e^t - 1)\mathbf{q}^T(\mathbf{I} - e^t\mathbf{Q})^{-1}\mathbf{1} + 1. \quad (\text{A3.8})$$

Hence, without showing the calculations, we follow the steps in Kritzinger (2011, p. 215), with the first moment being equal to

$$\begin{aligned} E(RL) &= M'(0) = e^0\mathbf{q}^T(\mathbf{I} - e^0\mathbf{Q})^{-1}\mathbf{1} + (e^{2(0)} - e^0)\mathbf{q}^T(\mathbf{I} - e^0\mathbf{Q})^{-2}\mathbf{Q}\mathbf{1} \\ &= \mathbf{q}^T(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}. \end{aligned} \quad (\text{A3.9})$$

Following this, the second moment is equal to

$$\begin{aligned}
E(RL^2) &= M''(0) = e^0 \mathbf{q}^T (\mathbf{I} - e^0 \mathbf{Q})^{-1} \mathbf{1} + e^{2(0)} \mathbf{q}^T (\mathbf{I} - e^0 \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} \\
&\quad + (2e^{2(0)} - e^0) \mathbf{q}^T (\mathbf{I} - e^0 \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} \\
&\quad + 2(e^{3(0)} - e^{2(0)}) \mathbf{q}^T (\mathbf{I} - e^0 \mathbf{Q})^{-3} \mathbf{Q}^2 \mathbf{1}. \\
&= \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} + \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} + \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1}. \tag{A3.10}
\end{aligned}$$

Equation (A3.6)	Equation (A3.7)
$ \begin{aligned} Var(RL) &= E(RL^2) - E(RL)^2 \\ &= \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} + \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} + \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} \\ &\quad - (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1})^2 \\ &= \mathbf{q}^T (\mathbf{I} + (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{Q} + (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{Q}) (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} \\ &\quad - (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1})^2 \\ &= \mathbf{q}^T ((\mathbf{I} - \mathbf{Q}) + \mathbf{Q} + \mathbf{Q}) (\mathbf{I} - \mathbf{Q})^{-1} (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} \\ &\quad - (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1})^2 \\ &= \mathbf{q}^T (\mathbf{I} + \mathbf{Q}) (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{1} - (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1})^2. \end{aligned} $	$ \begin{aligned} Var(RL) &= E(RL^2) - E(RL)^2 \\ &= \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} + \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} + \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} \\ &\quad - (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1})^2 \\ &= (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} + \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1}) \\ &\quad - (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1})^2 + (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}) \\ &= 2\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} - (ARL)^2 + ARL \end{aligned} $
$\therefore SDRL = \sqrt{\mathbf{q}^T (\mathbf{I} + \mathbf{Q}) (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{1} - (\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1})^2}$	$\therefore SDRL = \sqrt{2\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} - (ARL)^2 + ARL}$

Therefore Equations (A3.6) and (A3.7) are the same.

Proof 3.4 - Equations 3.7 & 3.8

Scariano and Calzada (2009) proposed alternative formulas to represent Equations (3.5) and (3.6). These formulas are preferred by authors in the literature because the *ARL* and the *SDRL* can be calculated without involving a matrix inversion.

Let $\beta(\delta)$ denote the probability that an \bar{X} sub-chart will not mark a sample as nonconforming on the first subsequent sample after a shift in the process mean from μ_0 to $\mu_0 \pm \delta\sigma_0$ with $\delta \neq 0$. The detection power of the \bar{X} sub-chart is given by $\theta = 1 - \beta(\delta)$. Calzada and Scariano (2001) noted that since the \bar{X} and the CRL sub-charts work in tandem to signal an OOC status, we then need proceed as follows in order to calculate the *ARL* and the *SDRL*:

Let Y_i denote the number of samples observed until the i^{th} nonconforming sample mean is signalled from the \bar{X} sub-chart. Let $Y_0 \equiv 0$, it follows that $(Y_1 - Y_0)$ follows a geometric

distribution with parameter θ . Define Y_2, Y_3, \dots similarly and note that the increments $(Y_1 - Y_0), (Y_2 - Y_1), (Y_3 - Y_2), \dots$ are mutually independent, each having a geometric distribution with parameter θ and an expected value of

$$E(Y_{i+1} - Y_i) = \frac{1}{\theta}. \quad (\text{A3.11})$$

As defined in Section 1.10, N denotes the first index for which a signal is given in the CRL sub-chart with LCL denoted by H . That is, $N \sim \text{geo}(P[Y_1 - Y_0 \leq H]) = \text{geo}(1 - (1 - \theta)^H)$. The first and second moments of N are given by

$$E(N) = \frac{1}{1 - (1 - \theta)^H} \quad (\text{A3.12})$$

and

$$E(N^2) = \frac{2 - (1 - (1 - \theta)^H)}{(1 - (1 - \theta)^H)^2} = \frac{1 + (1 - \theta)^H}{(1 - (1 - \theta)^H)^2}. \quad (\text{A3.13})$$

Moreover, let M denote the run-length variable of some basic sub-chart. The run-length is known to have a geometric distribution with a probability of marking a sample as nonconforming given by θ . Thus, the pmf, first and second moments of the random variable M are given by

$$P(M = w) = \theta(1 - \theta)^{w-1} \quad (\text{A3.14})$$

$$E(M) = \frac{1}{\theta} \quad (\text{A3.15})$$

$$E(M^2) = \frac{2 - \theta}{\theta^2}. \quad (\text{A3.16})$$

respectively. Note,

$$P[(Y_l - Y_{l-1}) = w] = P[M = w] \quad (\text{A3.17})$$

$$P[(Y_l - Y_{l-1})] = E[M] \quad (\text{A3.18})$$

$$P[(Y_l - Y_{l-1})^2] = E[M^2]. \quad (\text{A3.19})$$

Since the synthetic chart signal at the first index, say l , for which $Y_l - Y_{l-1} \leq H$, the increments $Y_i - Y_{i-1}$ and N are not mutually independent. Thus the conditional pmf of $Y_{i+1} - Y_i$ given $N = l$ is used to find the average and standard deviation of the run-length.

The *ARL* of the synthetic chart is given by

$$\begin{aligned} ARL &= E(Y_N) = E(E[Y_l | N = l]) = E\left(E\left[\sum_{i=1}^l (Y_i - Y_{i-1})\right] \middle| N = l\right) \\ &= E\left((N - 1) \frac{E(N)}{E(N - 1)} \left[\sum_{w>H} w \times P(M = w)\right]\right) + E(N) \times \left[\sum_{w=1}^H w \times P(M = w)\right] \\ &= E(N) \left[\sum_{w>H} w \times P(M = w)\right] + E(N) \times \left[\sum_{w=1}^H w \times P(M = w)\right] \\ &= E(N) \times \left[\sum_{w=1}^{\infty} w \times P(M = w)\right] \\ &= E(N) \times E(M) \end{aligned}$$

Therefore from Equations (A3.12) and (A3.15) it follows that

$$ARL = \frac{1}{\theta} \times \frac{1}{1 - (1 - \theta)^H}. \quad (\text{A3.20})$$

The *SDRL* of the synthetic chart is calculated as follows.

$$\begin{aligned}
\text{Var}(Y_N) &= E[\text{Var}(Y_l|N = l)] + \text{Var}[E(Y_l|N = l)] \\
&= E \left[\text{Var} \left(\sum_{i=1}^{l-1} (Y_i - Y_{i-1}) + (Y_l - Y_{l-1}) | N = l \right) \right] + \text{Var} \left[E \left(\sum_{i=1}^{l-1} (Y_i - Y_{i-1}) + (Y_l - Y_{l-1}) | N = l \right) \right] \\
&= \text{Var}[Y_1 - Y_0 | M > H] \times E(N - 1) + \text{Var}[Y_N - Y_{N-1} | M \leq H] + E^2[Y_1 - Y_0 | M > H] \times \text{Var}(N - 1) \\
&= \left\{ \frac{E(N)}{E(N - 1)} \sum_{w=H+1}^{\infty} w^2 \times P(M = w) - \frac{E^2(N)}{E^2(N - 1)} \left(\sum_{w=H+1}^{\infty} w \times P(M = w) \right)^2 \right\} E(N - 1) + E(N) \left\{ \sum_{w=1}^H w^2 \times P(M = w) \right\} \\
&\quad - E^2(N) \left\{ \sum_{w=1}^H w \times P(M = w) \right\}^2 + \frac{E^2(N)}{E^2(N - 1)} \left\{ \sum_{w=H+1}^{\infty} w \times P(M = w) \right\}^2 E(N) \times E(N - 1) \\
&= E(N) \times E(M^2) - \frac{E^2(N)}{E(N - 1)} \left(\sum_{w=H+1}^{\infty} w \times P(M = w) \right)^2 - E^2(N) \left\{ \sum_{w=1}^H w \times P(M = w) \right\}^2 + \frac{E^3(N)}{E^2(N - 1)} \left\{ \sum_{w=H+1}^{\infty} w \times P(M = w) \right\}^2 \\
&= E(N) \times E(M^2) + E^2(N) \left\{ \left(\sum_{w=H+1}^{\infty} w \times P(M = w) \right)^2 - \left(\sum_{w=1}^H w \times P(M = w) \right)^2 \right\} \\
&= E(N) \times E(M^2) + E^2(N) E(M) \left\{ E(M) - 2 \sum_{w=1}^H w \times P(M = w) \right\}.
\end{aligned}$$

Hence, using Equations (A3.12) to (A3.16) it follows that

$$\begin{aligned}
\text{Var}(Y_N) &= \left(\frac{1}{1 - (1 - \theta)^H} \right) \left(\frac{2 - \theta}{\theta^2} \right) + \left(\frac{1}{1 - (1 - \theta)^H} \right)^2 \left(\frac{1}{\theta} \right) \left\{ \frac{1}{\theta} - 2 \sum_{w=1}^H w \times \theta (1 - \theta)^{w-1} \right\} \\
&= \frac{2 - \theta}{\theta^2 (1 - (1 - \theta)^H)} + \frac{1}{(1 - (1 - \theta)^H)^2} \times \left\{ \left(\frac{1}{\theta} \right) \times \left(\frac{1}{\theta} - 2 \theta \sum_{w=1}^H w \times (1 - \theta)^{w-1} \right) \right\} \\
&= \frac{2 - \theta}{\theta^2 (1 - (1 - \theta)^H)} + \frac{1}{(1 - (1 - \theta)^H)^2} \times \left\{ \frac{1}{\theta^2} - 2 \sum_{w=1}^H w \times (1 - \theta)^{w-1} \right\}.
\end{aligned}$$

Therefore, the *SDRL* is given by

$$\text{SDRL} = \sqrt{ \frac{2 - \theta}{\theta^2 (1 - (1 - \theta)^H)} + \frac{1}{(1 - (1 - \theta)^H)^2} \times \left\{ \frac{1}{\theta^2} - 2 \sum_{w=1}^H w \times (1 - \theta)^{w-1} \right\} }. \quad (\text{A3.21})$$

3.7.2 Appendix 3B: SAS® programs

- **Optimal search algorithm to determine k and H**

(i) Zero-state mode

In Table 3.2, we illustrate this algorithm for $\delta = 0.75$, $n = 5$ and ARL_0 equal to 370.4 is easily done using Microsoft® Excel:

Step 1. Specify $\mu = 0$, $\sigma = 1$, $n = 5$, $\delta = 0.75$.

Step 2. Assume on the previous iteration $H = 1$ and we obtained $ARL_1(0.75) = 6.40581$.
Now, take $H = 2$.

Step 3. Find k by solving

$$\frac{1}{2\Phi(-k)} \times \frac{1}{1 - [1 - 2\Phi(-k)]^2} = 370.4$$

The corresponding value of k when $H = 2$ is 2.085 (using the Solver tool on Microsoft® Excel).

Step 4. First calculate $\theta = \theta(0.75) = 1 - [\Phi(2.085 - 0.75\sqrt{5}) - \Phi(-2.085 - 0.75\sqrt{5})]$ and $ARL_1(0.75) = \frac{1}{\theta} \times \frac{1}{1 - (1 - \theta)^2} = 5.16177$.

Step 5. Since $ARL_1(0.75)$ is equal to 6.40581 when $H = 1$ and when $H = 2$ it is equal to 5.16177, then according to the algorithm we go back to Step 3. We continue with these iterations until increasing H no longer decreases the corresponding $ARL_1(0.75)$. In a situation, we then proceed to Step 6.

(ii) Steady-state mode

```
proc iml;

ARL_0=370.4;
mu=0; stdev=1;
shift=0; n=1;

do H=1 to 10;
Q=J(H+1,H+1,0);
a1=J(1,4,99999);
one=J(H+1,1,1);
initial_vec=J(H+1,1,0);
initial_vec[2]=1; *head-start feature;
G=J(H+1,H+1,0);
G[1,]=1; G[1,1]=2;

do j=2 to H+1;
G[j,j]=1;
end;
u=J(H+1,1,0);
u[1,1]=1;
I=I(H+1);
*print G, u, I, initial_vec;

do k=1 to 3 by 0.0001;
theta=1-cdf("Normal",k-shift*sqrt(n),mu,stdev)+cdf("Normal",-k-
    shift*sqrt(n),mu,stdev);
Q[1,1]=1-theta;
Q[1,2]=theta;
Q[H+1,1]=1-theta;
do jj=2 to H;
Q[jj,jj+1]=1-theta;
end;
inv=inv(I-Q);

q1=inv(G-(Q)`)*u;
s=inv(one)`*q1)*q1;
*print G, u, I, Q, inv, one, s;
ARL_vec=inv*one;
ARL=s`*ARL_vec;*Using Markov chain (see Fu and Lou (2003));
diff=abs(ARL-ARL_0);
a=k||theta||ARL||diff;
a1=a1//a;
k__theta__ARL__diff=a1[2:nrow(a1),];
call sort(k__theta__ARL__diff,{4});
optimal_K=k__theta__ARL__diff[1,1];
ICARL=k__theta__ARL__diff[1,3];
end;

print H optimal_K ICARL;
end;
run;
quit;
```

- **Zero-state ARL calculation**

A direct calculation using Equation (3.2) in Microsoft® Excel is much easier to use.

- **Steady-state ARL calculation**

We use the following SAS® program to calculate the ARL values:

```

proc iml;
H=6;
k=2.2714; n=5;
ARL_0=370.4;
mu=0; stdev=1; shift=0.75;
a1=J(1,2,99999);
*do shift=0 to 3 by .1;
Q=J(H+1,H+1,0);
one=J(H+1,1,1);
initial_vec=J(H+1,1,0);
initial_vec[2]=1; *head-start feature;
G=J(H+1,H+1,0);
G[1,]=1; G[1,1]=2;

do j=2 to H+1;
G[j,j]=1;
end;
u=J(H+1,1,0);
u[1,1]=1;
I=I(H+1);
theta=1-cdf("Normal",k-shift*sqrt(n),mu,stdev)+cdf("Normal",-k-
shift*sqrt(n),mu,stdev);
Q[1,1]=1-theta;
Q[1,2]=theta;
Q[H+1,1]=1-theta;
do jj=2 to H;
Q[jj,jj+1]=1-theta;
end;
inv=inv(I-Q);
q1=inv(G-(Q`))*u;
s=inv(one`*q1)*q1;
ARL_vec=inv*one;
ARL=s`*ARL_vec;*Using Markov chain (see Fu and Lou (2003));
a=shift||ARL;
a1=a1//a;
shift__ARL=a1[2:nrow(a1),];
*end;
print shift__ARL;
run;
quit;

```

3.7.3 Appendix 3C: Comparison of four variables charts to monitor the process mean

Introduction

There is a vast number of control charts that have been proposed in the literature to monitor the process mean. In this section, we compare the performance of four different types of control charts to monitor the process mean for variables data. The four control charts are (i) the Shewhart \bar{X} chart, (ii) the synthetic chart (see Wu and Spedding (2000)), (iii) the 2-of-2 KL runs-rule chart (see Klein (2000)), and (iv) the 2-of-3 KL runs-rule chart (see Klein (2000)). In this comparison, it is assumed that the quality characteristics are normally distributed with a known mean μ_0 and a known standard deviation σ_0 .

Markov chain approach

The transition probability matrix (TPM) of the Markov chain for any general (integer) value of $M > 0$ are given by

$$\mathbf{P}_{(M+1, M+1)} = \left(\begin{array}{c|c} \mathbf{Q}_{(M, M)} & \mathbf{r}_{(M, 1)} \\ \hline \mathbf{0}'_{(1, M)} & \mathbf{1}_{(1, 1)} \end{array} \right) \quad (\text{A3.22})$$

where $\mathbf{Q}_{(M, M)}$ is the matrix of transient probabilities, the vector $\mathbf{r}_{(M, 1)}$ satisfies $\mathbf{r} = \mathbf{1} - \mathbf{Q}\mathbf{1}$ with $\mathbf{1}_{(M, 1)} = (1 \ 1 \ \dots \ 1)^T$ and $\mathbf{0} = (0 \ 0 \ \dots \ 0)^T$. Fu and Lou (2003, p. 73) proved that the first moment of the run-length is given by $\boldsymbol{\xi}(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}$, where $\boldsymbol{\xi}$ is the initial probability vector, that is,

$$ARL = \boldsymbol{\xi}(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} \quad (\text{A3.23})$$

Furthermore, it can be shown that the second moment using the method in Fu and Lou (2003, Chapter 5) that the standard deviation of the run-length is equal to

$$SDRL = \sqrt{\boldsymbol{\xi}(\mathbf{I} + \mathbf{Q})(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{1} - (\boldsymbol{\xi}(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1})^2}. \quad (\text{A3.24})$$

Shewhart \bar{X} chart

The upper and lower control limits of the Shewhart \bar{X} chart are given by $UCL_{\bar{X}}/LCL_{\bar{X}} = \mu_0 \pm k \frac{\sigma_0}{\sqrt{n}}$. Since it is customary to use $k = 3$, so that three-sigma limits are employed (see Montgomery (2013, p. 236)), thus we use this value to implement the \bar{X} chart.

The ARL is given by

$$ARL = \frac{1}{1 - \Phi(k - \delta\sqrt{n}) + \Phi(-k - \delta\sqrt{n})} \quad (A3.25)$$

where $\delta = \frac{|\mu_1 - \mu_0|}{\sigma_0}$ denotes the value of the shift in the process mean, where μ_1 is the OOC mean.

Synthetic chart

In Table A3.1 we give the optimal values of H and k using the algorithm in Wu and Spedding (2000) for the zero-state and steady-state modes. Davis and Woodall (2002, see p. 202 & p. 204) showed that using the same optimal values of H and k in the zero-state and in the steady-state modes one obtains significantly different performance results for the synthetic chart, implying that results depend on which mode of analysis is assumed. Moreover, in Table A3.1 we see that the values of k in zero-state and steady-state are different, with the steady-state values of k being lower than those in zero-state. Thus, it is our opinion that Davis and Woodall (2002, see p. 202) used approximately correct values for k in their Table 1, however, in their Table 2 (see Davis and Woodall (2002, p. 204)) the values of k are incorrect. This is further supported by the fact that in Davis and Woodall (2002, p. 204)'s Table 2, all the values of the IC ARL are much higher than 370.4. The correct values of k for the steady-state mode are given here in Table A3.1 obtained using a SAS® program given in the Appendix 3B.

Table A3.1: Optimal values of H and k with the corresponding values of in-control ARL for the zero-state and steady-state modes for a nominal ARL of 370.4

H	Zero-state		Steady-state	
	k	IC ARL	k	IC ARL
1	1.9435	370.4	1.9328	370.3
2	2.0848	370.4	2.0706	370.4
3	2.1640	370.4	2.1472	370.5
4	2.2188	370.4	2.1997	370.3
5	2.2604	370.4	2.2395	370.4
6	2.2939	370.4	2.2714	370.4
7	2.3218	370.4	2.2978	370.3
8	2.3458	370.4	2.3204	370.3
9	2.3667	370.4	2.3401	370.4
10	2.3852	370.4	2.3575	370.5
20	2.5032	370.4	2.4666	370.4
30	2.5690	370.4	2.5261	370.4
40	2.6142	370.4	2.5663	370.4
50	2.6483	370.4	2.5963	370.4

2-of-2 KL runs-rule chart

The 2-of-2 KL runs-rule chart was proposed in Klein (2000) where a control chart gives an OOC signal when either two successive points plot above an UCL , or two successive points plot below a LCL . Splitting a control chart into three regions i.e. one above the UCL (upper region), one below the LCL (lower region) and one between the control limits (center region). Denote the corresponding probability of a point falling in each region as pU , pL and p , respectively. Thus using the Markov chain approach, the transient states matrix is given by

$$\mathbf{Q} = \begin{bmatrix} p & pU & pL \\ p & 0 & pL \\ p & pU & 0 \end{bmatrix}.$$

Note state {1} (column 1) in the above matrix indicate no points beyond either of the control limits, state {2} (column 2) a point above UCL and state {3} (column 3) a point below LCL . Klein (2000) showed that the optimal value of k is equal to 1.781419. The above probabilities are given by $pU = 1 - \Phi(k - \delta\sqrt{n})$, $pL = 1 - \Phi(-k - \delta\sqrt{n})$ and $p = 1 - pU - pL$. The ARL is given by Expression (2) with $k = 1.781419$ and $\xi = (1 \ 0 \ 0 \ \dots \ 0)$.

2-of-3 KL runs-rule chart

The 2-of-3 KL runs-rule chart was proposed in Klein (2000) where a control chart gives an OOC signal when either two of three successive points plot above an UCL , or two of three successive points plot below a LCL . Similarly, using the Markov chain approach, the transient states matrix is given by

$$\mathbf{Q} = \begin{bmatrix} p & pU & pL & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & pL & p & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p & pU \\ 0 & 0 & 0 & 0 & 0 & p & 0 \\ p & 0 & pL & 0 & 0 & 0 & 0 \\ p & pU & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p & 0 \end{bmatrix}.$$

Note state {1} (column 1) in the above matrix indicate two successive points between the control limits, state {2} (column 2) a first point between control limits and the second point above UCL , state {3} (column 3) a first point between control limits and the second point above LCL , state {4} (column 4) a first point above the UCL and the second below the LCL , state {5} (column 5) a first point above the UCL and the second between control limits, state {6} (column 6) a first point above the LCL and the second between control limits, state {7} (column 7) a first point above the LCL and the second point above the UCL . Klein (2000) showed that the optimal value of k is equal to 1.930701. The ARL is given by Expression (2) with $k = 1.930701$ and $\xi = (1 \ 0 \ 0 \ \dots \ 0)$.

Discussion

In Table A3.2 we observe that the 2-of-3 KL chart performs better than all the other charts for small shifts of size $\delta \leq 0.6$. Moreover, the 2-of-3 KL chart always performs equally as or better than the 2-of-2 KL chart and equally as or better than the Shewhart \bar{X} chart except when $\delta > 2.7$. The 2-of-2 KL chart better than the Shewhart \bar{X} chart and the zero-state synthetic chart when $\delta \leq 0.4$. Also the Shewhart \bar{X} chart performs equally as the zero-state synthetic chart when $\delta = 5$. Finally, for moderate to large shifts (i.e. $\delta > 0.6$), the zero-state synthetic chart performs better than all the other charts. In addition, as the shift

increase from medium to large, the optimal values of H decrease from a large value to a small value.

However, in Table A3.3, the *2-of-3* KL chart performs better than all the other charts for small to medium shifts of size $\delta \leq 1.7$ (that is, a large degree of shifts compared to zero-state). Similarly, the *2-of-2* KL chart also now performs better than the Shewhart \bar{X} chart and the steady-state synthetic chart for a larger degree of shifts than in zero-state (i.e. $\delta \leq 1.3$). Moreover, the Shewhart \bar{X} chart performs better than all the other charts for shifts of size δ equal to 4 and 5. Finally, for the steady-state, the synthetic chart only performs better than the other charts for shifts of size $1.7 \leq \delta \leq 3$.

With respect to the deterioration in the performance of the synthetic chart from the zero-state to steady-state mode, this was originally observed by Davis and Woodall (2002) and has been well documented by a number of authors in the literature. Moreover, it goes without saying that the *2-of-3* KL chart performs well for small shifts and this chart is recommended if such shifts are of interest. Whereas, for moderate to large shift the synthetic chart yields better results than the competing charts discussed here.

Table A3.2: Comparison of the *ARL* values for the zero-state synthetic chart (for various values of *H* and *k*), $3\sigma \bar{X}$ chart, 2-of-2 KL chart and the 2-of-3 KL chart

	<i>H</i>	1	2	3	4	5	6	7	8	9	10	20	30	40	50	3σ limits	2-of-2 KL	2-of-3 KL
	<i>k</i>	1.9435	2.0848	2.1640	2.2188	2.2604	2.2939	2.3218	2.3458	2.3667	2.3852	2.5032	2.5690	2.6142	2.6483			
Shift	0	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	370.4	372.7
	0.1	354.2	352.3	351.3	350.5	350.0	349.5	349.1	348.8	348.5	348.3	346.9	346.1	345.7	345.4	352.9	342.0	341.6
	0.2	311.5	305.3	301.8	299.5	297.7	296.3	295.1	294.1	293.3	292.5	288.1	286.0	284.8	284.1	308.4	276.7	271.6
	0.3	255.4	245.1	239.5	235.7	232.9	230.7	228.9	227.4	226.1	225.0	218.5	215.6	214.0	213.1	253.1	207.1	199.5
	0.4	198.9	186.3	179.7	175.3	172.1	169.6	167.6	165.9	164.5	163.3	156.5	153.7	152.3	151.6	200.1	150.3	142.3
	0.5	149.8	136.9	130.3	126.1	123.0	120.6	118.8	117.2	115.9	114.8	109.0	106.8	105.9	105.6	155.2	108.5	101.4
	0.6	110.7	98.9	93.0	89.3	86.6	84.7	83.1	81.8	80.8	79.9	75.4	74.0	73.6	73.7	119.7	78.9	73.0
	0.7	81.2	71.1	66.2	63.1	61.0	59.4	58.2	57.2	56.4	55.7	52.6	51.8	51.9	52.3	92.3	58.2	53.4
	0.8	59.7	51.2	47.3	44.9	43.2	42.0	41.1	40.4	39.8	39.3	37.3	37.1	37.4	38.0	71.6	43.6	39.8
	0.9	44.1	37.3	34.1	32.3	31.0	30.1	29.4	28.9	28.5	28.2	27.0	27.2	27.7	28.4	55.8	33.3	30.2
	1	32.9	27.4	25.0	23.6	22.6	22.0	21.5	21.1	20.9	20.6	20.1	20.4	21.1	21.8	43.9	25.8	23.4
	1.1	24.8	20.5	18.6	17.5	16.8	16.3	16.0	15.8	15.6	15.5	15.3	15.8	16.5	17.2	34.8	20.3	18.4
	1.2	19.0	15.5	14.1	13.2	12.7	12.4	12.2	12.0	11.9	11.8	12.0	12.5	13.2	13.8	27.8	16.3	14.8
	1.3	14.7	11.9	10.8	10.2	9.8	9.6	9.5	9.4	9.3	9.3	9.6	10.2	10.8	11.4	22.4	13.2	12.0
	1.4	11.6	9.4	8.5	8.0	7.8	7.6	7.5	7.4	7.4	7.4	7.8	8.4	9.0	9.5	18.2	10.9	10.0
	1.5	9.2	7.4	6.8	6.4	6.2	6.1	6.1	6.1	6.1	6.1	6.5	7.1	7.6	8.0	15.0	9.2	8.4
	1.6	7.5	6.0	5.5	5.2	5.1	5.0	5.0	5.0	5.0	5.1	5.6	6.0	6.4	6.8	12.4	7.8	7.2
	1.7	6.1	4.9	4.5	4.3	4.3	4.2	4.2	4.2	4.3	4.3	4.8	5.2	5.5	5.8	10.3	6.7	6.2
	1.8	5.1	4.1	3.8	3.7	3.6	3.6	3.6	3.6	3.7	3.7	4.2	4.5	4.8	5.0	8.7	5.9	5.4
	1.9	4.3	3.5	3.2	3.1	3.1	3.1	3.1	3.2	3.2	3.3	3.7	4.0	4.2	4.4	7.4	5.2	4.8
	2	3.7	3.0	2.8	2.7	2.7	2.8	2.8	2.8	2.9	2.9	3.3	3.5	3.7	3.9	6.3	4.6	4.3
	2.1	3.2	2.6	2.5	2.4	2.4	2.5	2.5	2.5	2.6	2.6	2.9	3.1	3.3	3.4	5.4	4.2	3.9
	2.2	2.8	2.3	2.2	2.2	2.2	2.2	2.2	2.3	2.3	2.4	2.6	2.8	2.9	3.1	4.7	3.8	3.6
	2.3	2.4	2.1	2.0	2.0	2.0	2.0	2.1	2.1	2.1	2.1	2.4	2.5	2.7	2.7	4.1	3.5	3.3
	2.4	2.2	1.9	1.8	1.8	1.8	1.9	1.9	1.9	2.0	2.0	2.2	2.3	2.4	2.5	3.6	3.2	3.1
	2.5	2.0	1.7	1.7	1.7	1.7	1.7	1.7	1.8	1.8	1.8	2.0	2.1	2.2	2.3	3.2	3.0	2.9
	2.6	1.8	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.7	1.7	1.7	1.9	2.0	2.1	2.9	2.8	2.8
	2.7	1.7	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.6	1.6	1.6	1.7	1.8	1.9	2.6	2.7	2.6
2.8	1.5	1.4	1.4	1.4	1.4	1.4	1.4	1.5	1.5	1.5	1.5	1.6	1.7	1.7	1.8	2.4	2.6	2.5
2.9	1.4	1.3	1.3	1.3	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.5	1.6	1.6	1.7	2.2	2.5	2.4
3	1.4	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.4	1.4	1.4	1.4	1.5	1.5	1.6	2.0	2.4	2.4
4	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.2	2.0	2.0
5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.0	2.0

Table A3.3: Comparison of the *ARL* values for the steady-state synthetic chart (for various values of *H* and *k*), $3\sigma \bar{X}$ chart, 2-of-2 KL chart and 2-of-3 KL chart

	<i>H</i>	1	2	3	4	5	6	7	8	9	10	20	30	40	50	3σ limits	2-of-2 KL	2-of-3 KL	
	<i>k</i>	1.9328	2.0706	2.1472	2.1997	2.2395	2.2714	2.2978	2.3204	2.3401	2.3575	2.4666	2.5261	2.5663	2.5963				
Shift	0	370.3	370.4	370.5	370.3	370.4	370.4	370.3	370.4	370.4	370.5	370.4	370.4	370.4	370.4	370.4	370.4	372.7	
	0.1	354.7	353.2	352.3	351.6	351.2	350.9	350.6	350.4	350.2	350.1	349.3	348.9	348.7	348.7	352.9	342.0	341.6	
	0.2	313.2	307.9	305.1	303.2	301.8	300.9	300.0	299.3	298.8	298.4	295.8	294.9	294.5	294.4	308.4	276.7	271.6	
	0.3	258.5	249.7	245.2	242.1	240.1	238.5	237.2	236.2	235.4	234.7	231.1	229.9	229.5	229.5	253.1	207.1	199.5	
	0.4	203.2	192.4	187.0	183.6	181.2	179.5	178.0	176.9	176.1	175.3	171.7	170.7	170.6	170.7	200.1	150.3	142.3	
	0.5	154.7	143.7	138.4	135.1	132.8	131.2	129.9	128.9	128.2	127.5	124.6	124.0	124.2	124.6	155.2	108.5	101.4	
	0.6	115.8	105.7	101.0	98.2	96.3	94.9	93.9	93.1	92.5	92.0	90.0	89.8	90.2	90.7	119.7	78.9	73.0	
	0.7	86.2	77.6	73.7	71.4	69.8	68.8	68.0	67.4	67.0	66.6	65.4	65.5	66.1	66.8	92.3	58.2	53.4	
	0.8	64.3	57.2	54.0	52.2	51.1	50.3	49.7	49.3	49.0	48.8	48.1	48.5	49.2	49.9	71.6	43.6	39.8	
	0.9	48.3	42.5	40.1	38.7	37.9	37.3	36.9	36.6	36.4	36.3	36.0	36.6	37.3	38.0	55.8	33.3	30.2	
	1	36.7	32.1	30.2	29.2	28.5	28.1	27.8	27.6	27.5	27.4	27.5	28.1	28.8	29.5	43.9	25.8	23.4	
	1.1	28.2	24.5	23.1	22.3	21.8	21.5	21.4	21.2	21.2	21.1	21.4	22.0	22.7	23.4	34.8	20.3	18.4	
	1.2	22.0	19.0	17.9	17.3	17.0	16.8	16.7	16.6	16.6	16.5	16.9	17.6	18.3	18.9	27.8	16.3	14.8	
	1.3	17.3	15.0	14.1	13.7	13.5	13.3	13.2	13.2	13.2	13.2	13.7	14.3	15.0	15.6	22.4	13.2	12.0	
	1.4	13.9	12.0	11.3	11.0	10.8	10.8	10.7	10.7	10.7	10.7	11.2	11.9	12.5	13.0	18.2	10.9	10.0	
	1.5	11.3	9.8	9.2	9.0	8.9	8.8	8.8	8.8	8.8	8.8	9.4	10.0	10.5	11.0	15.0	9.2	8.4	
	1.6	9.3	8.1	7.6	7.5	7.4	7.3	7.3	7.4	7.4	7.4	8.0	8.5	9.0	9.4	12.4	7.8	7.2	
	1.7	7.7	6.7	6.4	6.3	6.2	6.2	6.2	6.2	6.3	6.3	6.9	7.4	7.8	8.1	10.3	6.7	6.2	
	1.8	6.5	5.7	5.5	5.4	5.3	5.3	5.3	5.4	5.4	5.5	6.0	6.4	6.8	7.0	8.7	5.9	5.4	
	1.9	5.6	4.9	4.7	4.6	4.6	4.6	4.7	4.7	4.7	4.8	5.3	5.6	5.9	6.2	7.4	5.2	4.8	
	2	4.8	4.3	4.1	4.1	4.1	4.1	4.1	4.1	4.2	4.2	4.2	4.7	5.0	5.3	5.4	6.3	4.6	4.3
	2.1	4.2	3.8	3.6	3.6	3.6	3.6	3.7	3.7	3.7	3.8	3.8	4.2	4.5	4.7	4.8	5.4	4.2	3.9
	2.2	3.8	3.4	3.3	3.2	3.3	3.3	3.3	3.3	3.4	3.4	3.5	3.8	4.0	4.2	4.3	4.7	3.8	3.6
	2.3	3.4	3.0	2.9	2.9	3.0	3.0	3.0	3.1	3.1	3.2	3.5	3.7	3.8	3.9	4.1	3.5	3.5	3.3
	2.4	3.0	2.8	2.7	2.7	2.7	2.7	2.8	2.8	2.8	2.9	2.9	3.2	3.3	3.5	3.6	3.6	3.2	3.1
2.5	2.8	2.5	2.5	2.5	2.5	2.5	2.6	2.6	2.6	2.7	2.7	2.9	3.1	3.2	3.2	3.2	3.0	2.9	
2.6	2.6	2.3	2.3	2.3	2.4	2.4	2.4	2.4	2.5	2.5	2.5	2.7	2.8	2.9	3.0	2.9	2.8	2.8	
2.7	2.4	2.2	2.2	2.2	2.2	2.2	2.3	2.3	2.3	2.3	2.4	2.5	2.6	2.7	2.8	2.6	2.7	2.6	
2.8	2.2	2.1	2.1	2.1	2.1	2.1	2.1	2.2	2.2	2.2	2.2	2.4	2.5	2.5	2.6	2.4	2.6	2.5	
2.9	2.1	2.0	2.0	2.0	2.0	2.0	2.0	2.1	2.1	2.1	2.1	2.2	2.3	2.4	2.4	2.2	2.5	2.4	
3	2.0	1.9	1.9	1.9	1.9	1.9	2.0	2.0	2.0	2.0	2.0	2.1	2.2	2.2	2.3	2.0	2.4	2.4	
4	1.6	1.5	1.5	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.2	2.0	2.0		
5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.0	2.0	2.0		

Conclusion

In this section, we compared four variables control charts (the Shewhart \bar{X} chart, the 2-of-2 KL chart, the 2-of-3 KL chart and the synthetic chart) to monitor the process mean. It was observed that the 2-of-3 KL performs better than all the other charts for very small shifts and the 2-of-2 KL chart also performs better than the other charts for small shifts, except the 2-of-3 KL chart which always perform better than or equally to 2-of-2 KL chart. Further, we observed that the zero-state synthetic chart has a better performance than the steady-state chart. Moreover, both the zero-state and steady-state synthetic chart have better performance for moderate to large shifts. Also, the findings here, invalidate the findings in Wu and Spedding (2000, Figure 1) that the synthetic chart always perform better than the runs-rule chart.

A topic that would interesting to pursue is to investigate the performance comparison between the 2-of-3 KL runs-rule chart and the EWMA and CUSUM charts that Davis and Woodall (2002) show perform better than the synthetic chart.

3.7.4 Appendix 3D: SAS® programs for the runs-rule charts

(i) 2-of-2 KL ARL values

```
proc iml;
k=1.781419;
mu=0; stdev=1; n=1;
a1=J(1,4,99999);
do shift=0 to 3 by .1;
pL=CDF("Normal",-k-shift*sqrt(n),mu,stdev);
pU=1-CDF("Normal",k-shift*sqrt(n),mu,stdev);
p=1-pU-pL;
ARL1=1/(1-p-(pU/(1+pU))-(pL/(1+pL)));*see Eq.(A1) on p.430 in Klein (2000);
*=====Markov Chain Approach=====*;

Q=J(3,3,0);
Q[,1]=p;
Q[1,2]=pU;
Q[3,2]=pU;
Q[1,3]=pL;
Q[2,3]=pL;
I=I(3);
inv=inv(I-Q);
one=J(3,1,1);
initial_vec={1 0 0};
ARL_vec=inv*one;
ARL2=initial_vec*ARL_vec; *Using Markov chain (see Fu and Lou (2003));
*****;
M2=initial_vec*(I+Q)*(ginv((I-Q)**2))*one;
* Calculating the standard deviation;
SDRL=sqrt(M2-((ARL2)**2));
a=shift||ARL1||ARL2||SDRL;
a1=a1//a;
shift__ARL=a1[2:nrow(a1),];
end;
*print Q, I, inv, one;
print shift__ARL [label = 'Shift__ARL1__ARL2__SDRL' format=.1];
run;
quit;
```

(ii) 2-of-3 KL ARL values

```
proc iml;
k=1.930701;
mu=0; stdev=1; n=1;
a1=J(1,3,99999);
do shift=0 to 3 by .1;
pL=CDF("Normal",-k-shift*sqrt(n),mu,stdev);
pU=1-CDF("Normal",k-shift*sqrt(n),mu,stdev);
p=1-pU-pL;
Q=J(7,7,0);
Q[1,1]=p; Q[5,1]=p; Q[6,1]=p; Q[2,5]=p; Q[7,5]=p; Q[3,6]=p; Q[4,6]=p;
Q[1,2]=pU; Q[6,2]=pU; Q[3,7]=pU;
Q[1,3]=pL; Q[2,4]=pL; Q[5,3]=pL;
I=I(7);
inv=inv(I-Q);
one=J(7,1,1);
initial_vec={1 0 0 0 0 0 0};
ARL_vec=inv*one;
ARL=initial_vec*ARL_vec;*Using Markov chain (see Fu and Lou (2003));
M2=initial_vec*(I+Q)*(ginv((I-Q)**2))*one;
* Calculating the standard deviation;
SDRL=sqrt(M2-((ARL)**2));
a=shift||ARL||SDRL;
a1=a1//a;
shift__ARL=a1[2:nrow(a1),];
end;
print shift__ARL [label = 'Shift__ARL__SDRL' format=.1];
run;
quit;
```

Chapter 4

Modified improved probability limits (MIPL) design for the synthetic Shewhart-type attributes charts

4.1 Introduction

There have been improvements in the field of SPC in developing more efficient designs for attribute control charts. For a thorough account of attributes control charts, see Woodall (1997) and more recently Szarka and Woodall (2011). In this chapter, we focus on uni-attribute control charts; for multiattribute charts see the review given by Topalidou and Psarakis (2009). Wu et al. (2001) introduced a synthetic control chart for attributes data, more specifically, the synthetic np chart which is a combination of a Shewhart np chart and the CRL chart. The reader is referred to Section 1.10 on a discussion of the synthetic chart and the CRL chart. Recall that, in designing a synthetic chart, it is assumed that a CRL sub-chart only has a LCL denoted by H . Also recall that (see Section 3.2) there are two methods that are widely used to compute the run-length characteristics of a synthetic charts, i.e. the seven steps “direct” approach given in Wu and Spedding (2000a) and the Markov chain approach proposed in Davis and Woodall (2002). However, recently Wu et al. (2010) and Haridy et al. (2012) presented a non-Markov chain approach to evaluate the run-length properties of a synthetic X chart and a synthetic np chart, respectively.

The concept of synthetic charts was introduced in Wu and Spedding (2000a) to monitor a shift in the mean. Since then, there have been a number of authors that have made significant contributions to this concept. For attributes charts, (i) Wu et al. (2001) proposed a synthetic version of the np chart, (ii) Bourke (2008) implemented the Markov chain approach given in Davis and Woodall (2002) to re-evaluate the performance of the synthetic np chart in detecting increases in fraction nonconforming, (iii) Castagliola et al. (2013) investigated the effect of parameter estimation for the synthetic p , np , c and u charts and more recently, (iv) Chong et al. (2014) studied the synthetic np chart with double sampling scheme.

In this chapter, we propose a modified improved probability limits (MIPL) method for the synthetic p , np , c and u charts. It should be noted that we remark on the design of the synthetic np and u charts and focus on the synthetic p and c charts. The concept of an improved probability limits (IPL) chart design was originally proposed by Zhang et al. (2004) for a standard geometric sub-chart. In this work, we modify and refine this approach and apply it to design what are called the MIPL synthetic p , np , c and u charts that result in better $AFAR$ values. In addition, these new and improved charts can be formulated such that they have the same or better OOC performance than the two traditional methods (i.e. k -sigma limits (k -SL) and conventional probability limits (CPL)). Therefore, the aim of this chapter is to extend on the contributions of Castagliola et al. (2013) in the case where parameters are known (Case K).

The rest of the chapter is structured as follows. The run-length properties used for evaluating the statistical performance of the methods for the synthetic p and c charts are discussed in Section 4.2. A discussion of the MIPL method for the synthetic p chart is given in Section 4.3 and the corresponding discussion for the synthetic c chart is done in Section 4.5; a review for the k -SL and CPL methods are given at the beginning of the respective sections. In Sections 4.4 and 4.6 examples and empirical comparisons among the three methods are done, providing an insight concerning the optimal design for the synthetic p and c charts, respectively. Concluding remarks are given in Section 4.7.

4.2 Properties of synthetic p and c control charts

Let (LCL_p, UCL_p) and (LCL_c, UCL_c) denote a pair of control limits (i.e. lower and upper) of the p sub-chart and the c sub-chart, respectively. In Section 3.2, we illustrated how the synthetic chart operates; discussed the two states under which the analysis of the run-length is done. Moreover, we demonstrated the equivalence of the synthetic control chart and the runs-rules chart with a head-start feature that motivated Davis and Woodall (2002) to formulate the Markov chain approach. Then, using the Markov chain approach, we presented and discussed the run-length properties. Furthermore, we showed and demonstrated the algorithm to choose the optimal values of k and H .

To this end, we assume that both the IC proportion nonconforming (p_0) and the IC number of nonconformities (c_0) are known or specified.

Run-length characteristics

The transition probability matrix that governs the Markov chain approach is given by Equation (3.2). So that the zero-state pmf and cdf are given by Equations (3.3) and (3.4), respectively. The simplified expression form of the zero-state ARL and $SDRL$ are given by

$$ARL = \frac{1}{\theta(1 - (1 - \theta)^H)} \quad (4.1)$$

and

$$SDRL = \sqrt{\frac{2 - \theta}{(1 - (1 - \theta)^H)\theta^2} + \frac{\frac{1}{\theta^2} - 2 \sum_{l=1}^H l(1 - \theta)^{l-1}}{(1 - (1 - \theta)^H)^2}} \quad (4.2)$$

respectively. When the process is IC, the FAR is typically used to design a chart in Case K. For attributes charts, the $AFAR$ of a particular chart (given by a set of control limits) may not necessarily be equal to the nominal value due to the discrete nature of distribution of the charting statistics. Thus,

$$AFAR = P(\text{Signal from a specific pair of control limits} | \text{IC}) = \theta(1 - (1 - \theta)^H). \quad (4.3)$$

The design of a control chart depends on the $AFAR$ value, the closer the $AFAR$ is to the nominal value (denoted by FAR_0), the better the chart is (or the chart design is). The nominal values of ARL and $SDRL$ are denoted by ARL_0 and $SDRL_0$, whereas the actual (or attained) values are denoted by $AARL_0$ and $ASDRL_0$, respectively. However, when the process is OOC, we denote these as $AARL_1$ and $ASDRL_1$, respectively.

Algorithm for the optimal values of (H, k)

The algorithm for the chart parameters (k and H) has already been discussed in Section 3.2.1. Furthermore, consider Table 4.1, we show the values of $AARL_1(\delta)$ for different combinations of k and H when δ is equal to 0.25, 0.75 and 1.5; and identify the pair of k and H that yields the smallest $AARL_1(\delta)$ for a sample of size 5, where $\mu = 0$ and $\sigma = 1$. The resulting optimal values are $(H, k) = (47, 2.639)$, $(7, 2.322)$ and $(2, 2.085)$, respectively, where the values in grey shading refer to the smallest $AARL_1(\delta)$.

Table 4.1: Values of the $AARL_1$ for different (H, k) combinations when $\delta = 0.25, 0.75$ and 1.5 for $n = 5$ with $ARL_0 = 370.4$

H	k	$AARL_1(0.25)$	H	k	$AARL_1(0.75)$	H	k	$AARL_1(1.5)$
1	1.943	125.46682	1	1.943	6.40581	1	1.943	1.17936
2	2.085	113.14676	2	2.085	5.16177	2	2.085	1.12554
3	2.164	106.92154	3	2.164	4.72298	3	2.164	1.13433
4	2.219	102.94004	4	2.219	4.52441	4	2.219	1.14725
5	2.260	100.10235	5	2.260	4.43126	5	2.260	1.15886
6	2.294	97.94800	6	2.294	4.39349			
7	2.322	96.24345	7	2.322	4.38795			
8	2.346	94.85472	8	2.346	4.40237			
9	2.367	93.69865	9	2.367	4.42966			
10	2.385	92.72005	10	2.385	4.46542			
11	2.402	91.88075						
12	2.417	91.15327						
13	2.431	90.51726						
14	2.443	89.95722						
15	2.455	89.46115						
16	2.466	89.01955						
17	2.476	88.62479						
18	2.486	88.27066						
19	2.495	87.95208						
20	2.503	87.66477						
21	2.511	87.40517						
22	2.519	87.17025						
23	2.526	86.95743						
24	2.533	86.76448						
25	2.540	86.58947						
26	2.546	86.43074						
27	2.552	86.28681						
28	2.558	86.15640						
29	2.564	86.03836						
30	2.569	85.93168						
31	2.574	85.83545						
32	2.579	85.74888						
33	2.584	85.67124						
34	2.589	85.60187						
35	2.593	85.54019						
36	2.598	85.48566						
37	2.602	85.43779						
38	2.606	85.39615						
39	2.610	85.36033						
40	2.614	85.32997						
41	2.618	85.30472						
42	2.622	85.28427						
43	2.625	85.26835						
44	2.629	85.25667						
45	2.632	85.24901						
46	2.636	85.24512						
47	2.639	85.24482						
48	2.642	85.24789						
49	2.645	85.25416						
50	2.648	85.26346						

Castagliola et al. (2013) implemented a search algorithm for (H, k) (similar to what is done in Table 4.1) for the synthetic \bar{X} chart with three distinct shifts of interest i.e. small ($\delta = 0.25$), moderate ($\delta = 0.75$) and large ($\delta = 1.5$) given $ARL_0 = 370.4$. Assuming a normal approximation to both the binomial and Poisson distributions, the synthetic p , np , c and u charts, that are expected to be optimal in detecting small, medium and large shifts, have the pairs (H, k) as given in Table 4.1. In this chapter, we follow the same approach as in Castagliola et al. (2013), since our objective is to improve the performance of the charts discussed therein. Therefore, constructing probability limits from the k -sigma limits, the nominal tail probability of the p or c sub-charts is given by

$$\tau = 2 \times (1 - \Phi(k)) . \quad (4.4)$$

So that the pairs $(H, k) = (47, 2.639)$, $(7, 2.322)$ and $(2, 2.085)$ may equivalently be written as $(H, \tau) = (47, 0.008315)$, $(7, 0.020233)$ and $(2, 0.037069)$, respectively.

Due to the discrete nature of the assumed distributions, these parameters (H, τ) do not guarantee an $AARL_0$ exactly equal to the nominal value of 370.4. Thus, our aim is to adopt (H, τ) and formulate a better method that ensures that, given the discrete nature of the binomial and Poisson distributions, we obtain an $AFAR$ that is as close as possible to the nominal values of the synthetic p and c charts and, when the process is OOC, it has nearly ARL -unbiased control limits.

Therefore, from this point onwards, we assume that the run-length properties are evaluated under the assumption of a zero-state mode. The steady-state performance will be reported elsewhere.

4.3 Statistical design of the synthetic p chart

Let $Y_{i,1}, \dots, Y_{i,n}$ for $i = 1, 2, \dots$ and $n \geq 1$ be a sample of independent random variables with $Y_i = \sum_{j=1}^n Y_{i,j} \sim \text{BIN}(p)$ where p is the proportion nonconforming units (when $p = p_0$, the process is IC). The probability θ that the p sub-chart will mark a sample as nonconforming is

$$\begin{aligned} \theta &= P(Y_i \leq nLCL_p | p) + P(Y_i \geq nUCL_p | p) && 0 < p < 1 \\ &= 1 - \begin{cases} \sum_{j=0}^b \binom{n}{j} p^j (1-p)^{n-j} & \text{if } nLCL_p < 0 \\ \sum_{j=0}^b \binom{n}{j} p^j (1-p)^{n-j} - \sum_{j=0}^a \binom{n}{j} p^j (1-p)^{n-j} & \text{if } nLCL_p \geq 0 \end{cases} && (4.5) \end{aligned}$$

The control limit constants, a and b , are related to the LCL and UCL , respectively, and are defined in Section 4.3.1.

4.3.1 Traditional methods of the synthetic p chart

In this section, we briefly summarize the traditional k -SL and CPL methods for Case K.

k -sigma limits

The control limits of the p sub-chart for the k -SL method are given in Section 1.9.1, Equation (1.11).

Conventional probability limits

The control limits of the p sub-chart for the CPL method are similar those given in Section 1.9.2, Equation (1.13), however the nominal tail area is different (see Equation (4.4)). Thus, for the synthetic p chart, the control limits are computed as follows. For the LCL_p , using we find the

largest positive integer $a = nLCL_p$ that makes the left tail probability, $P_l(Y_i \leq a|p_0)$, to be at most equal to $\frac{\tau}{2}$, i.e.

$$\sum_{j=0}^a \binom{n}{j} p_0^j (1 - p_0)^{n-j} \leq \frac{\tau}{2}. \quad (4.6)$$

Following this, $LCL_p = \frac{a}{n}$. If $a < 0$, it means that $LCL_p < 0$ and then we assume that the LCL_p does not exist since the proportion nonconforming is never negative. This arises when n and/or p_0 is small. In such a situation LCL_p is said to be not applicable (NA). For the UCL_p , we find the smallest positive integer $b = nUCL_p - 1$ that makes the right tail probability $P_r(Y_i \geq b + 1|p_0)$, to be at least equal to $1 - \frac{\tau}{2}$, i.e.

$$\sum_{j=b+1}^n \binom{n}{j} p_0^j (1 - p_0)^{n-j} \geq 1 - \frac{\tau}{2}. \quad (4.7)$$

Following this, $UCL_p = \frac{b+1}{n}$. In the event that LCL_p is not applicable, for Equation (4.7) we find $b = nUCL_p - 1$ so that $P_r(Y_i \geq b + 1|p_0) \geq 1 - \tau$.

Problems associated with traditional methods

As highlighted in Section 1.11.3, the traditional methods reviewed above have poor IC run-length characteristics, more so when n and/or p_0 are small. According to the recommendations in the literature, when $np_0(1 - p_0) > 5$ (i.e. when the central limit theorem approximation is good for the binomial distribution), these methods should yield attained run-length properties that are closer to their nominal values. For example, suppose that $FAR_0 = 0.0027$, $n = 100$ and $p_0 = 0.2$ so that $np_0(1 - p_0) = 20 > 5$. Consequently, we expect that the 3-SL and CPL methods will result in $AFAR$ values close to the FAR_0 . Suppose that $H = 2$ (see Table 4.1), for the 3-SL method, using Equations (1.11) and (4.5), we have $a = 11$ and $b = 28$ so that, from Equation (4.3), we obtain $AFAR = 0.00209$ which is 22.58% lower than the nominal value of 0.0027. On the other hand, for the CPL method, using Equations (4.6) and (4.7), we have $a = 11$ and $b = 29$ so that, from Equation (4.3), we obtain $AFAR = 0.00112$ which is 58.46%

lower than the nominal value of 0.0027. The corresponding $AARL_0$ values are 478.41 and 891.56, respectively. From this example it is clear that these traditional control charts yield $AFAR$ values that are significantly different from the FAR_0 , since for both methods, the charts will signal less often than what is nominally expected, especially for the CPL method.

In the next section we offer a solution to this problem by considering a chart design method called the MIPL. It will be shown that this method yields control limits that result in $AFAR$ and $ASDRL_0$ values that are the same or much closer to the nominal values compared to the two traditional methods. Moreover, we show that the MIPL method can be formulated such that it yields similar or better nearly ARL -unbiased control limits compared to the two traditional methods. First we consider the synthetic p chart.

4.3.2 Modified improved probability limits (MIPL) for the synthetic p chart

In Zhang et al. (2004), the IPL method for the geometric (sub-) chart was constructed such that the $AFAR \geq FAR_0$ (we call this an anti-conservative approach for the probability limits design). Here we modify and refine their method so that it takes into account both the conservative (i.e. $AFAR \leq FAR_0$) and the anti-conservative approach to designing probability limits (sub-) charts. Firstly, we generate set A , with all the values of $a = nLCL_p$, satisfying some condition that will be discussed below. Then for each $a \in A$, we find the pair (a, b_1) that yields a $\theta \leq \tau$ and similarly we find the pair (a, b_2) that yields a $\theta \geq \tau$. Hence, to this end, we let $T(a, b_i | n, p_0)$ for $i = 1, 2$ denote a set of control limit constants generated from set A , for some given p_0 . Next, we let T_1 denote a subset of T with control limit constants that yield $AFAR(a, b_1 | n, p_0) \leq FAR_0$ and similarly let T_2 denote a subset of T with control limit constants that yield $AFAR(a, b_2 | n, p_0) \geq FAR_0$. In this case, it is easy to see that $T_1 \cup T_2 = T$. Then proceed with the following steps.

Step 1: Generate set A .

Let $a = [nLCL_p] \in A = \{\text{NA}, 0, 1, \dots, L_{max}\}$ where L_{max} is equal to the largest integer such that

$$\sum_{j=0}^{L_{max}} \frac{e^{-c_0} c_0^j}{j!} \leq \tau \quad (4.8)$$

holds, for some n , p_0 and τ (as calculated in Equation (4.4)). Recall that “NA” stands for not applicable which implies that $LCL_c < 0$. Note the difference between Equations (4.8) and (4.6).

Step 2: For each $a \in A$, we find the two corresponding integer values of b such that

$$P_l(Y_i \leq a | c_0) + P_r(Y_i \geq b + 1 | c_0) = \theta \in T. \quad (4.9)$$

i.e. for each $a \in A$ in Step 1, we find $(a, b_i) \in T_i$ for $i = 1, 2$. Thus, in total we obtain $2 \times (L_{max} + 2)$ pairs of control limit constants.

Step 3: For each pair (a, b_i) in Step 2, we compute the $AFAR$, for some pre-calculated value of H using Equation (4.9) and then calculate the percentage relative deviation from the target FAR_0 defined by

$$D = 100 \times \left(\frac{AFAR(a, b_i, H | n, p_0) - FAR_0}{FAR_0} \right). \quad (4.10)$$

Step 4: Choose the pair $(a^*, b^*) \in T$ such that

$$AFAR(a^*, b^*, H | n, p_0) = \min_{(a,b) \in T} AFAR(a, b, H | n, p_0) \quad (4.11)$$

i.e. we choose the pair (a^*, b^*) that result in the minimum absolute deviation of $AFAR$ from the FAR_0 . The MIPL for the synthetic p chart are given by $LCL_p = a^*/n$ and $UCL_p = (b^* + 1)/n$.

To picture the MIPL procedure, assume that set H contains all possible pairs of control limit constants (a, b) where a and b are integers with $a < b$ and $a \in A$ (for the MIPL method) for some given p_0 . Some of these pairs will yield $AFAR$ values much closer to FAR_0 and others will yield $AFAR$ values that differ significantly from FAR_0 . Note that the control limit constants of all three methods (i.e. k -SL, CPL and MIPL) will be in set H . Furthermore, we let $T \subseteq H$ (see Equation (4.9)). That is, for each d , we find the corresponding two values of b (i.e. b_1 and b_2) such that the resulting $AFAR$ will either be at least equal to or at most equal to the target FAR_0 , respectively. Then, consequently the four step procedure for the MIPL method results in the pair (a^*, b^*) , which is as close as possible to the FAR_0 , that is, a local IC optimal pair in set T and, subsequently, this pair will be the global optimal pair also in set H .

Remark 1: Since the synthetic p chart is equivalent to the synthetic np chart when the sample size is fixed (see Castagliola et al. (2013)) it follows that the corresponding MIPL for the synthetic np chart are given by $LCL_{np} = a^*$ and $UCL_{np} = b^* + 1$. Recall that in this study we assume the FSS scheme and, consequently, the sample size is fixed.

In the next section, we give a numerical example and do an overall comparative study for the three methods considered here.

4.4 Illustrations and performance comparison of the synthetic p chart methods

We first give an illustrative example for all three methods and then we do an overall comparative performance study to see which method outperforms the rest.

4.4.1 Example

Example 4.1. Assume that a manufacturing production process operates at a proportion nonconforming (fallout) of 20% and that $n = 100$; with $FAR_0 = 0.0027$. In addition, we are interested in detection of large shifts (i.e. $\delta = 1.5$), hence $H = 2$ and $k = 2.085$ (see Table 4.1).

Chart Designs

The traditional charts for this example are constructed and illustrated in Section 4.3.1 and are displayed in Table 4.3. For the MIPL method, the calculations are as follows. From Step 1, the value of L_{max} that satisfies Equation (4.8) is equal to 12, so that $A = \{NA, 0, 1, \dots, 12\}$. For each $a \in A$, we find the corresponding b_i such that the $AFAR(a, b_i, H/100, 0.2)$ of these pairs is an element of T (see Step 2). Table 4.2 shows all the possible pairs of control limit constants in set T (with subsets T_1 and T_2 , as defined earlier), the $AFAR$ and the percentage relative deviation from the FAR_0 for each pair (calculated in Step 3). Then, using Step 4, the optimal values for the control limits constants of the MIPL method, when the process is IC, are obtained using Table 4.2 and are given by $(a^*, b^*) = (12, 29)$ with an $AFAR$ that is only 2.71% lower than 0.0027. This is indicated by the use of grey shading in Table 4.2.

Table 4.2: All pairs of control limits constants in set T for the synthetic p chart using the MIPL method for $n = 100$, $p_0 = 0.2$, $\tau = 0.037069$ and $FAR_0 = 0.0027$

Set T_1					Set T_2				
$a = nLCL_p$	$b = nUCL_p - 1$	θ	$AFAR$	$ D $	$a = nLCL_p$	$b = nUCL_p - 1$	θ	$AFAR$	$ D $
NA	27	0.03415	0.00229	15.08%	NA	26	0.05583	0.00606	124.46%
0	27	0.03415	0.00229	15.08%	0	26	0.05583	0.00606	124.46%
1	27	0.03415	0.00229	15.08%	1	26	0.05583	0.00606	124.46%
2	27	0.03415	0.00229	15.08%	2	26	0.05583	0.00606	124.46%
3	27	0.03415	0.00229	15.08%	3	26	0.05583	0.00606	124.46%
4	27	0.03416	0.00229	15.06%	4	26	0.05584	0.00606	124.49%
5	27	0.03417	0.00230	14.99%	5	26	0.05585	0.00606	124.61%
6	27	0.03423	0.00230	14.70%	6	26	0.05591	0.00608	125.08%
7	27	0.03443	0.00233	13.71%	7	26	0.05611	0.00612	126.66%
8	27	0.03501	0.00241	10.81%	8	26	0.05669	0.00624	131.29%
9	27	0.03649	0.00261	3.19%	9	26	0.05817	0.00657	143.33%
10	28	0.02572	0.00131	51.64%	10	27	0.03985	0.00311	15.28%
11	28	0.03260	0.00209	22.58%	11	27	0.04673	0.00426	57.95%
12	29	0.03658	0.00263	2.71%	12	28	0.04535	0.00402	48.88%

Table 4.3 summarizes the three charting methods along with the $AFAR$ values and the percentage relative deviation from the FAR_0 for each pair, the $AARL_0$ and the $ASDRL_0$ values, respectively.

Table 4.3: Comparison among three methods of the synthetic p chart for $n = 100$, $p_0 = 0.2$ and $FAR_0 = 0.0027$

Method	(a, b)	$AFAR$	$ D $	$AARL_0$	$ASDRL_0$
k -sigma limits	(11,28)	0.00209	22.58%	478.41	506.29
Conventional probability limits	(11,29)	0.00112	58.46%	891.56	930.68
Modified improved probability limits	(12,29)	0.00263	2.71%	380.67	405.23

It is clear that, for this example, the MIPL method results in control limits with $AARL_0$ and $ASDRL_0$ values that are almost close to the corresponding nominal values.

4.4.2 Empirical comparison of the synthetic p chart methods

In this section we perform an empirical comparative study for the three methods of the synthetic p chart (discussed in this Sections 4.3) assuming that the $FAR_0 = 0.0027$ (or, equivalently, $ARL_0 = 370.4$). In Figures 4.1 to 4.3 we present the graphical plots of the $AFAR$ and the corresponding $ASDRL_0$ values for the synthetic p chart for the following pairs $(H, k) = (47, 2.639)$, $(7, 2.322)$ and $(2, 2.085)$. Firstly, for these pairs, we notice that, as the $AFAR$ values converge to the $FAR_0 = 0.0027$, the corresponding $ASDRL_0$ values converge to approximately 457, 413 and 394, respectively. Note that this is in contrast to the ordinary p chart, where the $ASDRL_0$ values would have been expected to be approximately 370. Furthermore, for these pairs, the $AFAR$ and the $ASDRL_0$ values, which result from the k -SL method, are significantly different (i.e. high fluctuations) from the nominal values. The control limits of the CPL method yield values of the $AFAR$ ($ASDRL_0$) that are all lower than or equal to FAR_0 (greater than or equal to $SDRL_0$). The MIPL method does not perform well when the IC proportion nonconforming is small, but as the IC process parameter increases, the $AFAR$ and the $ASDRL_0$ values seem to fluctuate around the nominal values.

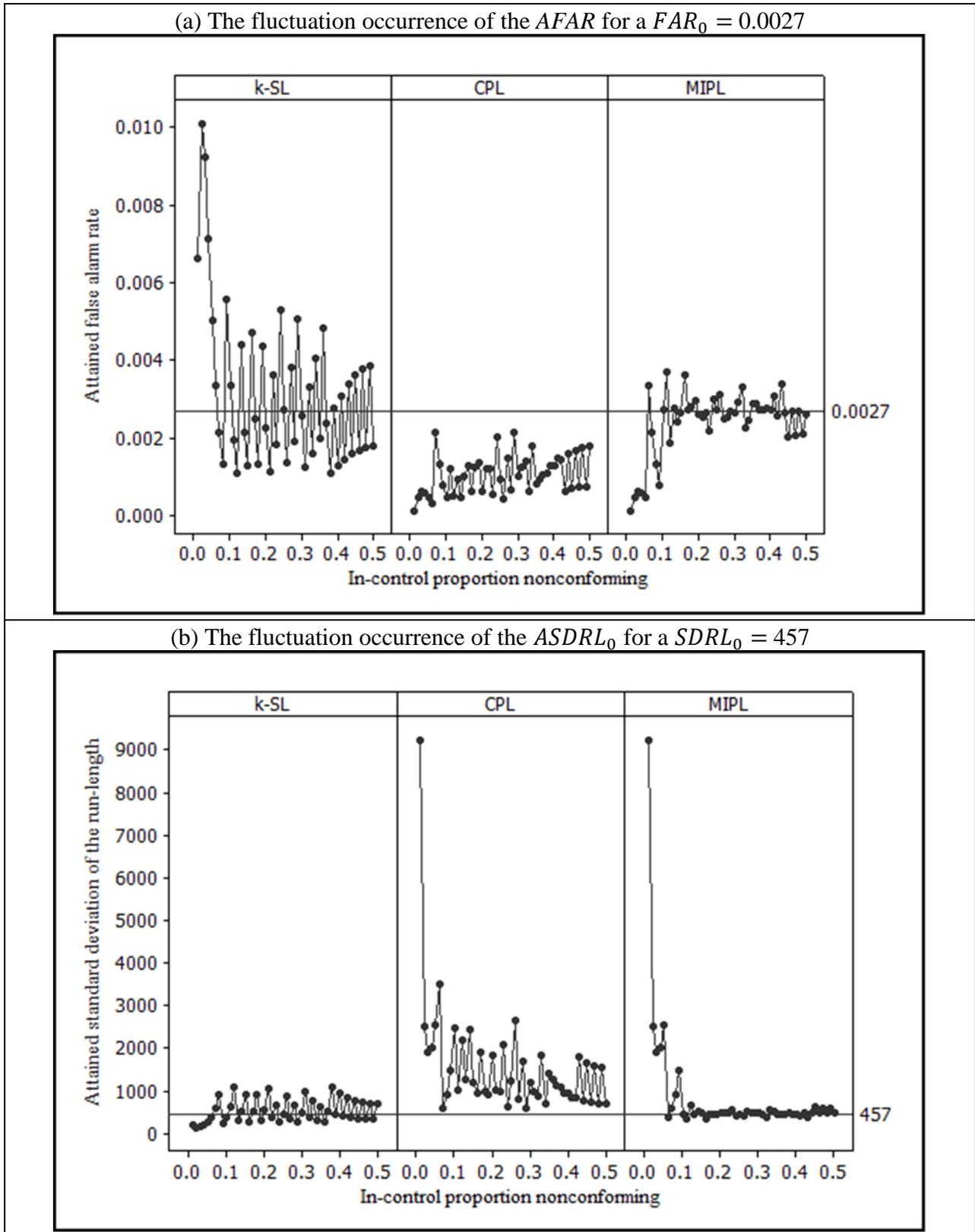


Figure 4.1: Comparison of the run-length characteristics among three methods of the synthetic p chart when the process is IC for $n = 50$ and varying $p_0 = 0.01(0.01)0.50$ for $(H = 47, k = 2.639)$

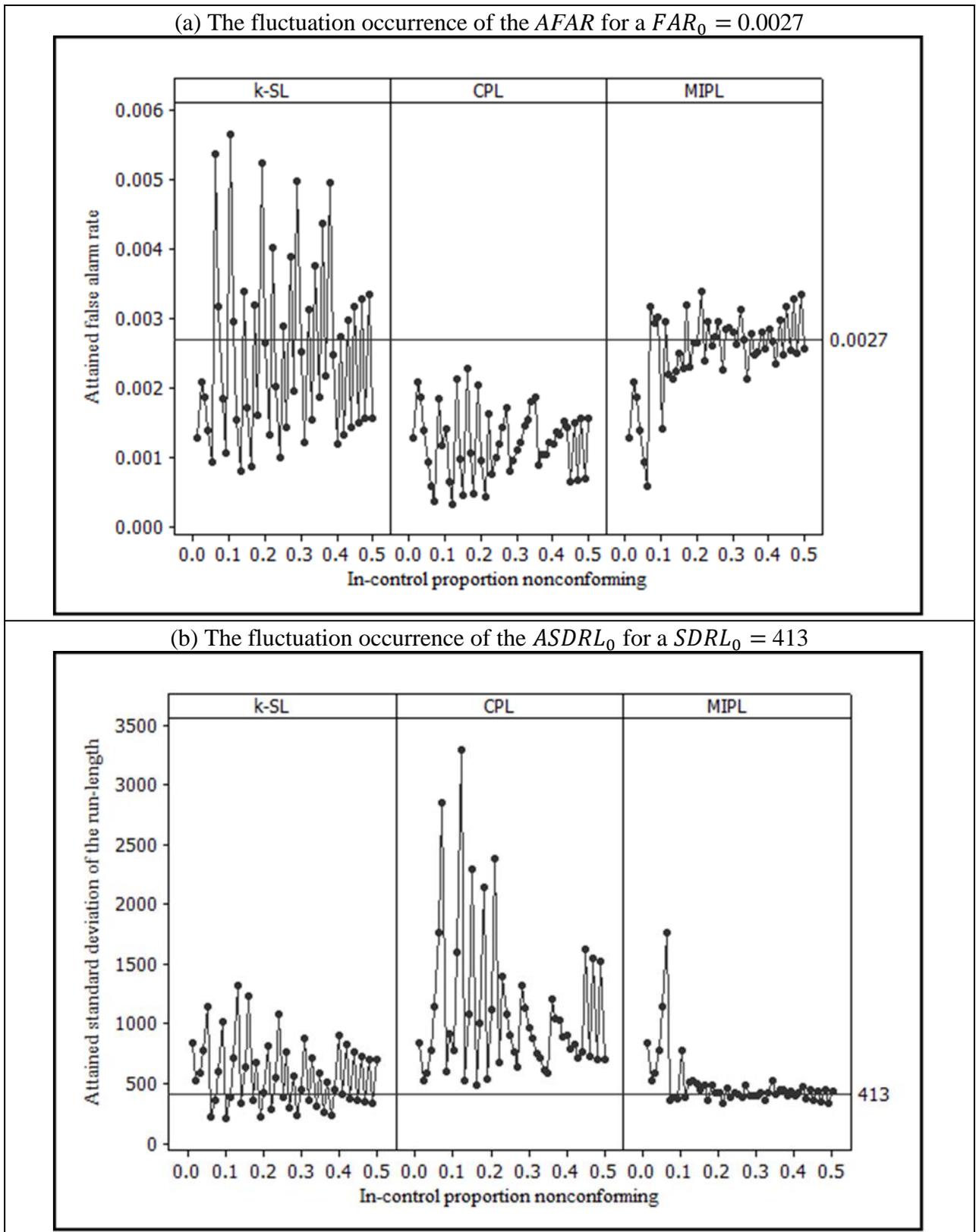


Figure 4.2: Comparison of the run-length characteristics among three methods of the synthetic p chart when the process is IC for $n = 50$ and varying $p_0 = 0.01(0.01)0.50$ for $(H = 7, k = 2.322)$

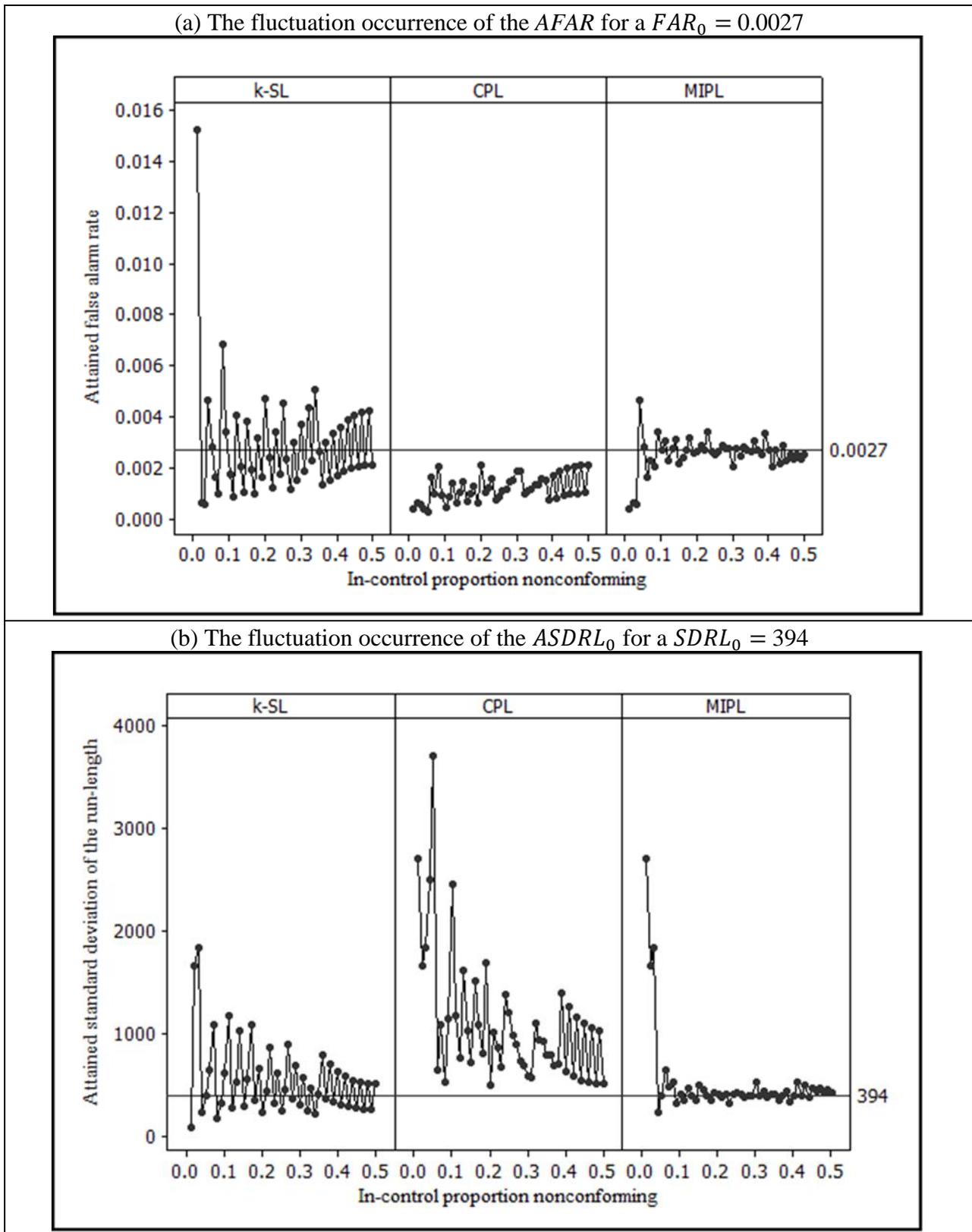
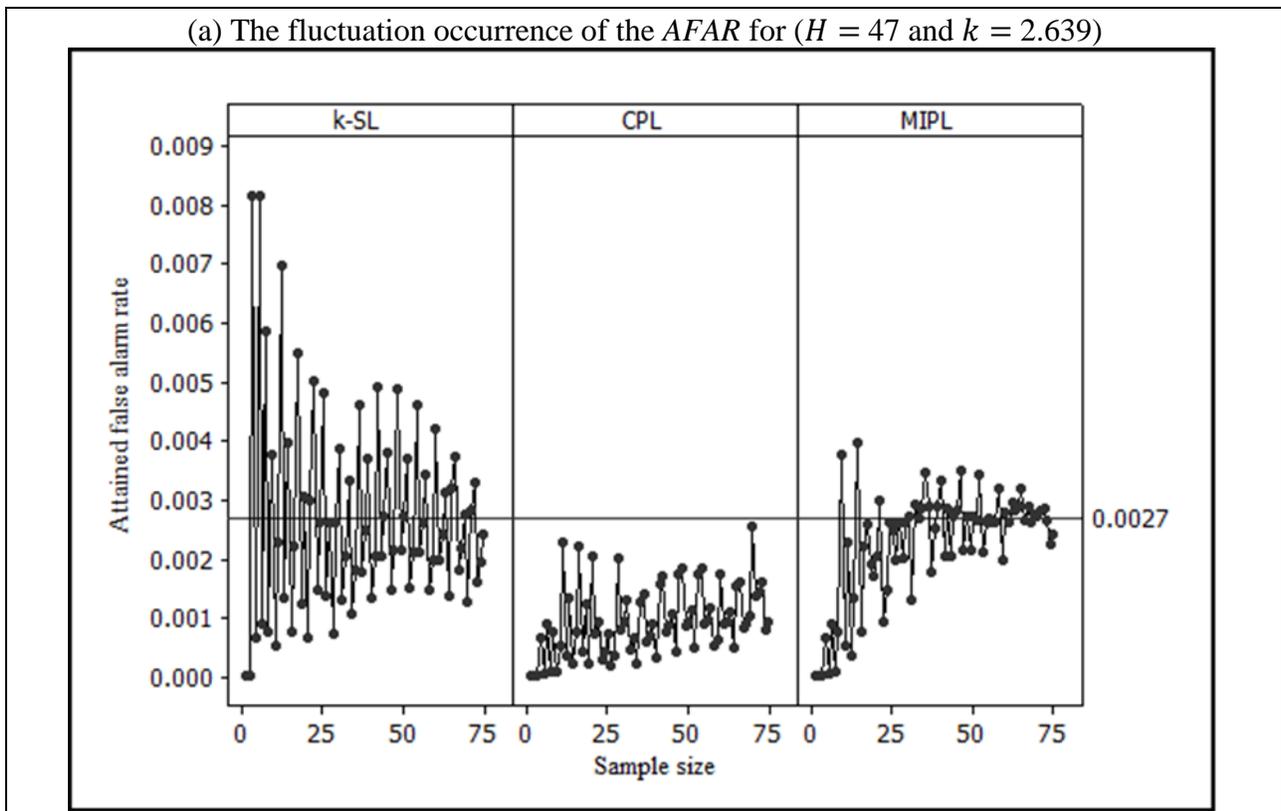
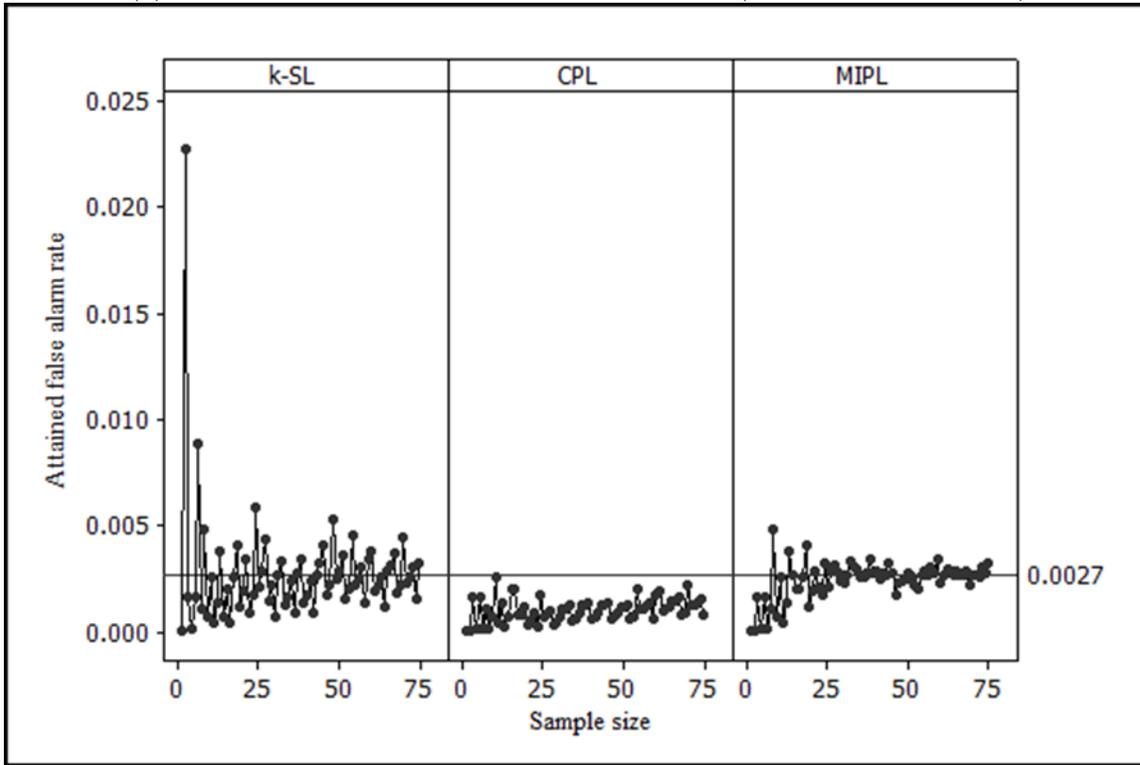


Figure 4.3: Comparison of the run-length characteristics among three methods of the synthetic p chart when the process is IC for $n = 50$ and varying $p_0 = 0.01(0.01)0.50$ for $(H = 2, k = 2.085)$

In Figure 4.4, we fix p_0 at 0.25 and the following can be observed. When the sample size is very small all the methods have undesirable values for the $AFAR$ and the $ASDRL_0$. However, for all three pairs of (H, k) , it can be seen that, as the sample size increases, the $AFAR$ values of the MIPL method tends towards the nominal value of 0.0027. This pattern (which is also observed for other combinations of n and p_0) indicates that, as the sample size increases, the MIPL method results in $AFAR$ values that tend towards the nominal value much faster than the two traditional methods. Note that the $AFAR$ values of the k -SL method tend towards the FAR_0 as a slower rate and the CPL method has been seen to be unreliable since it results in excessively high $ASDRL_0$ values (see Figures 4.1 to 4.3). Therefore, the MIPL method for the synthetic p chart results in better $AFAR$ and $ASDRL_0$ values compared to the k -SL and CPL methods.



(b) The fluctuation occurrence of the *AFAR* for ($H = 7$ and $k = 2.322$)



(c) The fluctuation occurrence of the *AFAR* for ($H = 2$ and $k = 2.085$)

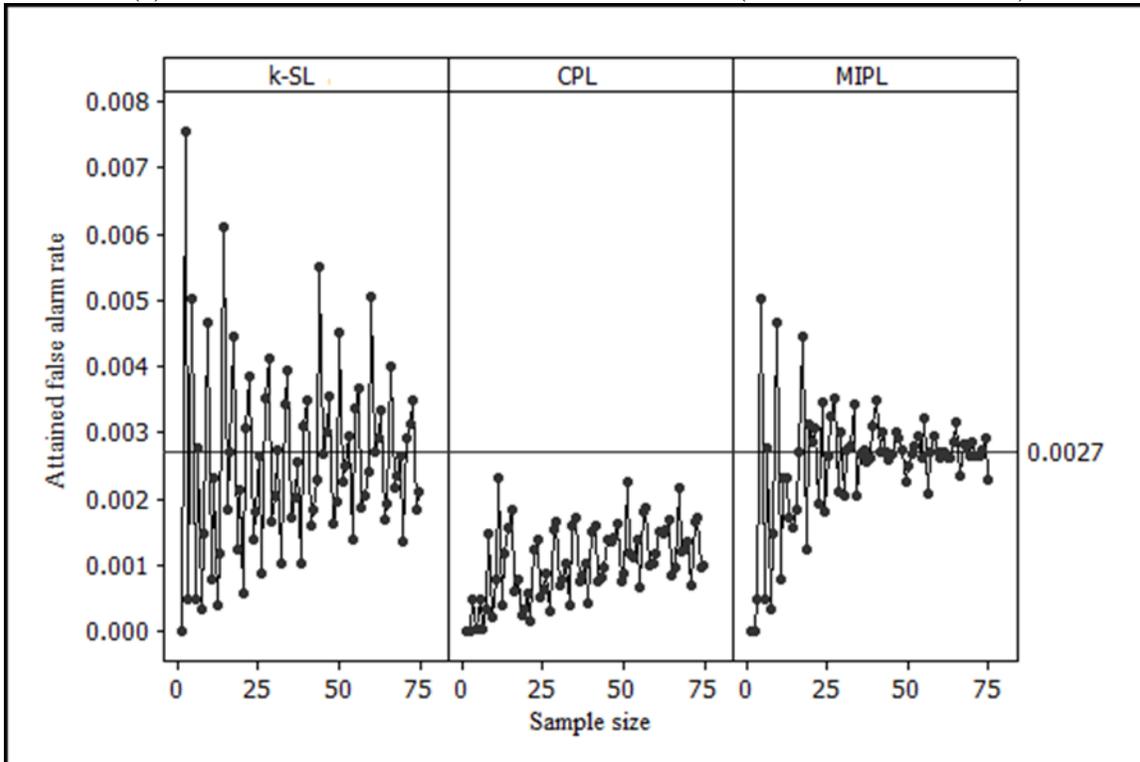


Figure 4.4: Comparison of the *AFAR* values for the method of the synthetic p chart when the process is IC for $p_0 = 0.25$ and varying $n = 1(1)50$

Therefore, Equation (4.11) ensures that the MIPL method yields $AFAR$ and $ASDRL_0$ values that are the same or much closer to the nominal values compared to the traditional methods. Although the MIPL method results in better $AFAR$ and $ASDRL_0$ values for the synthetic p chart; the most important function of a control chart is quick detection of a sustained process shift. Hence in the next section we illustrate how we can use the MIPL method to construct a synthetic p chart with better OOC performance than the two traditional methods.

4.4.3 Performance

It is generally known that, when the process is IC, the $AARL_0$ of a control chart should be large (preferably close to ARL_0) and when the process is OOC, the $AARL_1$ should be small. To study the OOC performance we use ARL curves (it shows an ARL for any possible value that parameter p can shift to (see Acosta-Mejia (1999))). For example, in Example 4.1 the control limit constants (12, 29) resulted in $AFAR$ and/or $AARL_0$ values much closer to FAR_0 and/or ARL_0 for the MIPL method, however, assuming that only sustained shifts with increments of 0.01 are of interest, this pair is not optimal in detecting small process deterioration, since the maximum of the ARL curve is not equal to $AARL_0$ (where $p_0 = 0.2$). That is, for some values of p ($\neq p_0$) $AARL_1 > AARL_0$. Control charts with this property were defined in Pignatiello et al. (1995) as ARL -biased charts. Acosta-Mejia (1999) showed that for the p chart, it is not always possible to obtain exact unbiasedness (because of the discrete nature of the charting statistics). Our aim is to construct control charts that are nearly ARL -unbiased using the MIPL method.

Step 1: We first let p^* be the value of the proportion nonconforming corresponding to the peak of the ARL curve, so that $ARL(a, b_i, H|n, p^*)$ is the value of the peak of the curve, with $(a, b_i) \in T$ defined in Equation (4.17). In addition, $ARL(a, b_i, H|n, p_0)$ is the $AARL_0$ when $p = p_0$, for some $i = 1, 2$.

Step 2: For each pair (a, b_i) in Step 1, we compute

$$q = ARL(a, b_i, H|n, p^*) - ARL(a, b_i, H|n, p_0). \quad (4.12)$$

Note that $q = 0$ if the synthetic p chart has ARL -unbiased control limit constants.

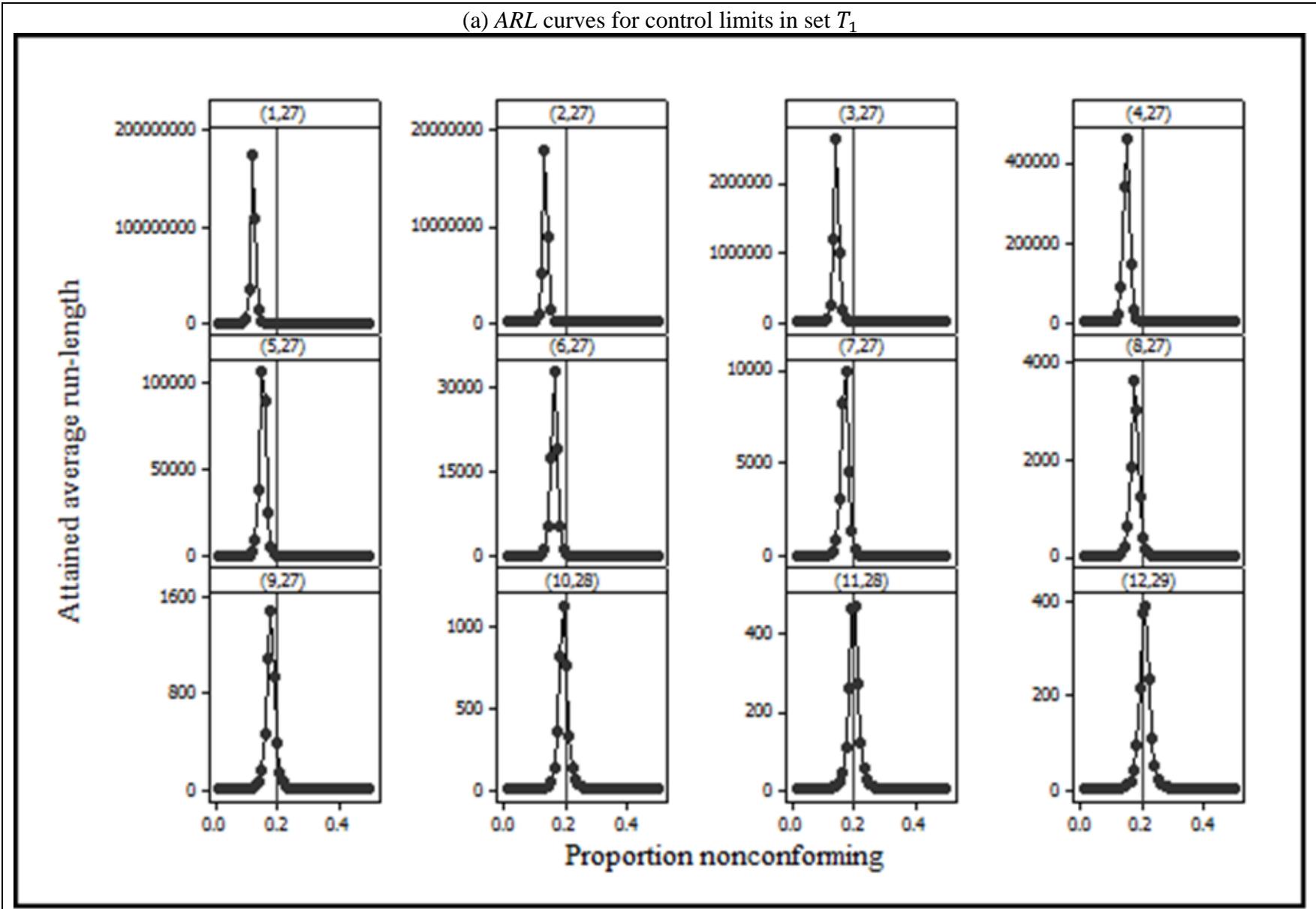
Step 3: Choose the pair $(a^\#, b^\#) \in T$ such that

$$ARL(a^\#, b^\#, H|n, p) = \min_{(a, b_i) \in T} ARL(a, b_i, H|n, p) \quad (4.13)$$

i.e. we choose the pair $(a^\#, b^\#)$ that result in the smallest value of q . Thus the nearly ARL -unbiased MIPL for the synthetic p chart are given by $LCL_p = a^\#/n$ and $UCL_p = (b^\# + 1)/n$.

To illustrate this, we use Example 4.1 to construct a synthetic p chart that will result in nearly ARL -unbiased control limits. Taking the pairs (a, b) in T that are given in Table 4.3 as the control limit constants, we construct the ARL curves (see Figures 4.5 (a) and (b)). It is evident that most of the control limit constants in set T have undesirable OOC values in addition to having a poor IC performance. The pairs (11, 28) and (12, 28) result in $q = 0$ (see Equation (4.15)), with an $AARL_0$ equal to 478.41 and 248.77, respectively. It seems more rational to choose the pair (11, 28) than (12, 28) since its $AARL_0$ is much closer to ARL_0 . Note that, when taking into account *both* the IC and OOC performance, the pair (12, 29) would be preferred over the latter two pairs which are exactly unbiased (see Figure 4.5).

(a) ARL curves for control limits in set T_1



(b) ARL curves for control limits in set T_2

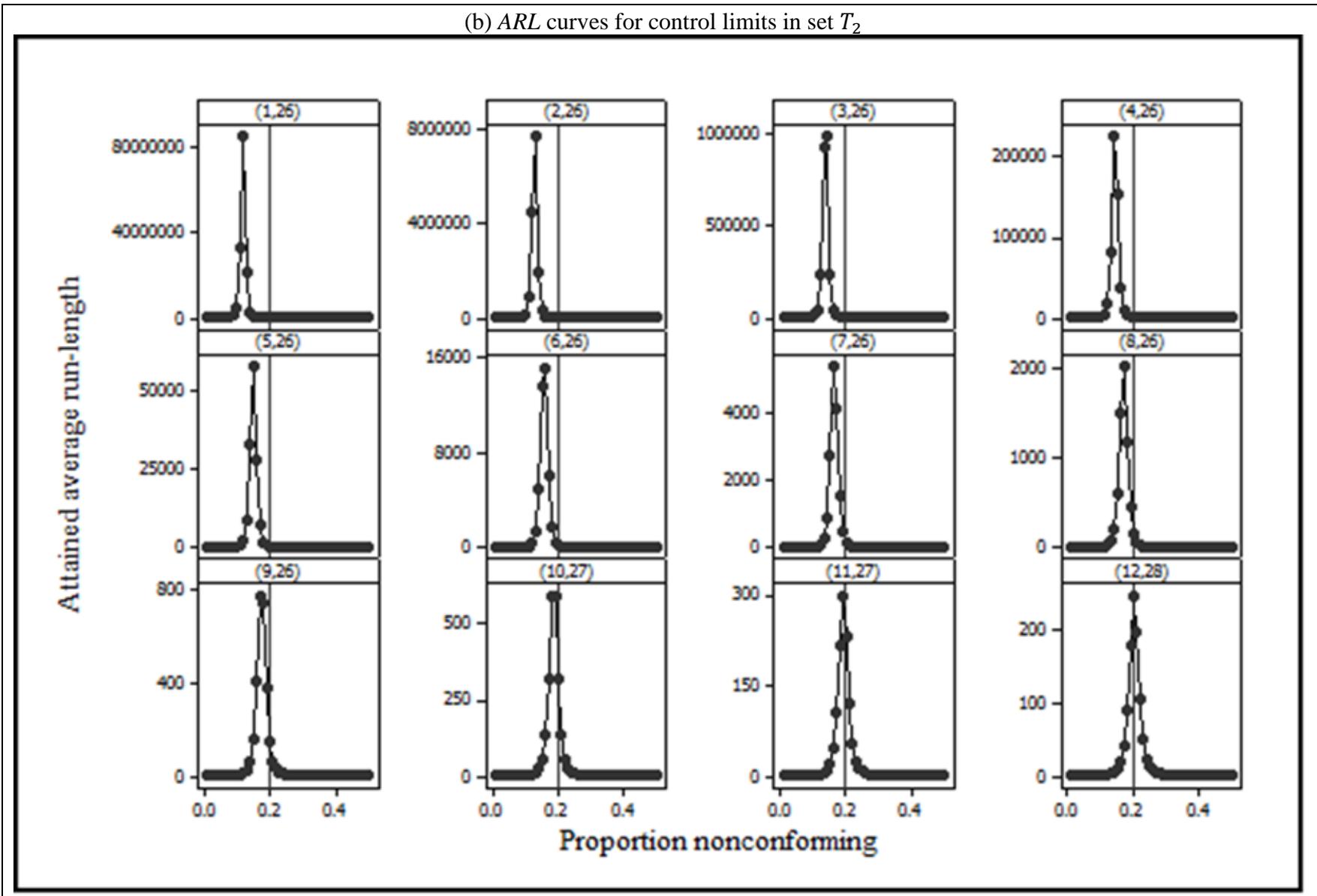


Figure 4.5: ARL curves of the control limits generated by the MIPL method for $p_0 = 0.2$ and $n = 100$ with $(H = 2, k = 2.085)$

In Figure 4.6, we plot the competing traditional methods' ARL curves along with the nearly ARL -unbiased MIPL pair (11, 28) and deduce that the k -SL method is ARL -unbiased with $AARL_0 = 478.41$. Furthermore, although the CPL method has ARL -unbiased control limits, the $AARL_0 = 891.56$ is much higher than 370.4. In this example, the k -SL and MIPL methods yield the same OOC performance, since they have the same control limits.

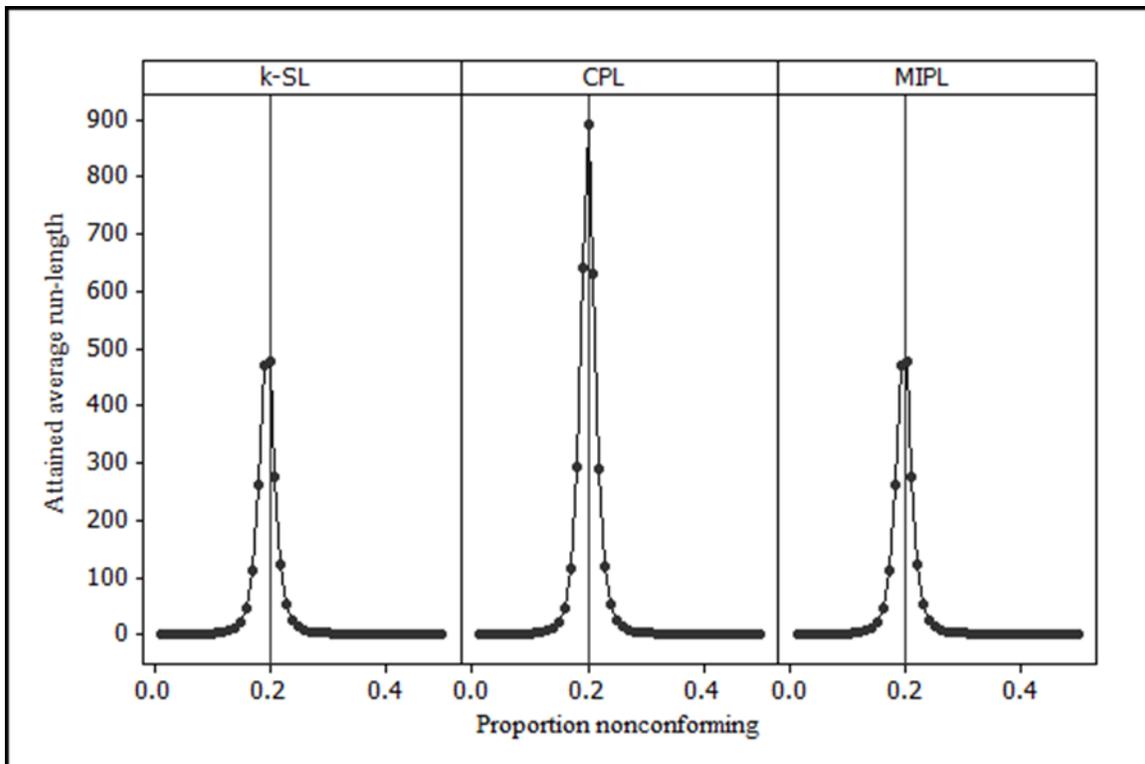


Figure 4.6: ARL curves for the three methods of the synthetic p chart with $p_0 = 0.2$ and $n = 100$

Simulations indicate that when the process parameters n and p_0 are both small, all three methods do not have a good OOC performance. In fact, for very small values of p_0 all three methods require a high value of n for the chart to be efficient. However, in most cases, the nearly ARL -unbiased MIPL method outperforms its competitors because the set T provides more options for the optimal pair $(a^\#, b^\#)$ compared to 3-SL and CPL methods, which only have one option for (a, b) .

Remark 2: Nearly *ARL*-unbiased control limits for the *np* chart

It follows similarly that the nearly *ARL*-unbiased control limits for the synthetic *np* chart are given by $LCL_{np} = a^\#$ and $UCL_{np} = b^\# + 1$.

4.5 Statistical design of the synthetic *c* chart

Let $Y_{i,1}, \dots, Y_{i,n}$, $i = 1, 2, \dots$ and $n \geq 1$ be a sample of independent random variables with $Y_i = \sum_{j=1}^n Y_{i,j} \sim POI(c)$ where c is the number of nonconformities (when $c = c_0$, the process is IC). The probability that the c sub-chart will mark an inspection unit as nonconforming is

$$\begin{aligned}\theta &= P(Y_i \leq LCL_c | c) + P(Y_i \geq UCL_c | c) && 0 < c < \infty \\ &= 1 + \sum_{j=0}^d \frac{e^{-c} c^j}{j!} - \sum_{j=0}^f \frac{e^{-c} c^j}{j!}\end{aligned}\tag{4.14}$$

4.5.1 Traditional methods for the synthetic *c* chart

k-sigma limits

The control limits of the c sub-chart for the Shewhart k -SL method is given in Section 1.9.1, see Equation (1.12).

Conventional probability limits

For the CPL method, the control limits are computed as follows. For the LCL_c , we find the largest integer $d = [LCL_c]$ that makes the left tail probability, $P_l(Y_i \leq d | c_0)$, to be at most equal to $\frac{\tau}{2}$. Thus,

$$\sum_{j=0}^d \frac{e^{-c_0} c_0^j}{j!} \leq \frac{\tau}{2} \quad (4.15)$$

So that $LCL_c = d$. Similarly, for the UCL_c we find the smallest integer f that makes the right tail probability $P_r(Y_i \geq f + 1 | c_0)$, to be at least equal to $1 - \frac{\tau}{2}$. Thus,

$$\sum_{j=f+1}^{\infty} \frac{e^{-c_0} c_0^j}{j!} \geq 1 - \frac{\tau}{2}. \quad (4.16)$$

In the event that LCL_p is not applicable, for Equation (4.16) we find f so that $P_r(Y_i \geq f + 1 | c_0) \geq 1 - \tau$.

Problem associated with the traditional methods

As indicated in Section 1.11.3, the traditional methods reviewed above have poor IC run-length properties, more especially when c_0 is small. However, when the normal approximation to the Poisson distribution is satisfied (i.e. $c_0 \geq 15$ see Montgomery (2013, p. 101)), one would expect that these methods should yield attained run-length properties that are close to the nominal values. For illustration, suppose that $FAR_0 = 0.0027$, $c_0 = 16$ and $H = 2$. For the 3-SL method, using Equations (1.12) and (4.14) we have $d = 7$ and $f = 24$ so that from Equation (4.3), yields $AFAR = 0.00205$ which is 23.90% lower than the nominal value of 0.0027. On the other hand, for the CPL method using (4.15) and (4.16) we have $a = 7$ and $b = 25$ so that from Equation (4.3) $AFAR = 0.00106$ which is 60.87% lower than the nominal value of 0.0027. The corresponding $AARL_0$ values are 486.66 and 946.47, respectively, and it is clear that these control charts are highly problematic for practical use, since the IC ARL is not close to the nominal value of 370.4.

In the next section, we offer a solution to this problem by considering a chart design called the MIPL. It will be shown that this method yields control limits that result in the $AFAR$ and the $ASDRL_0$ either the same as or much closer to the nominal values compared to the two

traditional methods. Moreover, the MIPL method can be formulated such that it yields similar or better nearly *ARL*-unbiased control limits compared to the traditional methods.

4.5.2 Modified improved probability limits (MIPL) for the synthetic c chart

Similar to the MIPL method for the synthetic p chart, we first generate set A . Further, we let $T(d, f_i | c_0)$ for $i = 1, 2$ denote a set of control limit constants generated from set A , for some given c_0 . Next, let T_1 denote a subset of T with control limits that yield an $AFAR(d, f_1 | c_0) \leq FAR_0$ and similarly let T_2 denote a subset of T with control limits that yield an $AFAR(d, f_2 | c_0) \geq FAR_0$. Then proceed with the following steps as in Zhang et al. (2004).

Step 1: Find all possible pairs of integers (d, f) in set T .

Let $d = LCL_c \in A = \{\text{NA}, 0, 1, \dots, L_{max}\}$ where L_{max} is equal to the largest integer such that

$$\sum_{j=0}^{L_{max}} \frac{e^{-c_0} c_0^j}{j!} \leq \tau \quad (4.17)$$

holds, for some τ and c_0 . “NA” stands for not applicable which implies that $LCL_c < 0$.

Step 2: For each $d \in A$, we find the corresponding integer value of f such that,

$$P_l(Y_i \leq d | c_0) + P_r(Y_i \geq f + 1 | c_0) = \theta \in T \quad (4.18)$$

i.e. for each $d \in A$ in Step 1, we find $(d, f_i) \in T_i$ for $i = 1, 2$. Thus, in total we obtain $2 \times (L_{max} + 2)$ pairs of control limit constants.

Step 3: Then for each pair (d, f_i) in Step 2, we compute the $AFAR$ for some pre-calculated value of H and then calculate the percentage relative deviation from the target FAR_0 defined by

$$D = 100 \times \left(\frac{AFAR(d, f_i, H|c_0) - FAR_0}{FAR_0} \right) \quad (4.19)$$

Step 4: Choose the pair $(d^*, f^*) \in T$ such that

$$AFAR(d^*, f^*, H|c_0) = \min_{(d, f_i) \in T} AFAR(d, f_i, H|c_0) \quad (4.20)$$

i.e. we choose the pair (d^*, f^*) that result in the minimum absolute deviation of $AFAR$ from the FAR_0 . The MIPL for the synthetic c chart are given by $LCL_c = d^*$ and $UCL_c = f^* + 1$.

Similarly to the MIPL procedure of the synthetic p chart in Section 4.3.2, the four step procedure for the MIPL method of the synthetic c chart results in the pair (d^*, f^*) , which is as close as possible to the target FAR_0 , that is, a local IC optimal pair in set T and subsequently, this pair will be the global optimal pair also in set T .

Remark 3: MIPL for the u chart

Similarly, the corresponding MIPL for the synthetic u chart can be formulated by assuming that $Y_{i,j} \sim POI(u)$ where u is the OOC average number of nonconformities per inspection unit, by defining $\bar{Y}_i = \frac{1}{n} \sum_{j=1}^n Y_{i,j}$ (see Castagliola et al. (2013)). For example, the Shewhart u sub-chart control limits are given in Montgomery (2013, p. 324) and using u_0 instead of c_0 in Equations (4.15) and (4.16) yield the corresponding CPL method.

In the next section, we consider a numerical example and comparative study for the three methods considered here.

4.6 Illustrations and performance comparison of the synthetic c chart methods

Firstly we illustrate the three methods using an example. Following this, an empirical comparison between the three methods are done by considering different values of the parameter c_0 . Furthermore, we show that the MIPL method can be formulated such that it yields similar or better nearly ARL -unbiased control limits compared to the traditional methods.

4.6.1 Example

Example 4.2. Assume that it is known that an inspection unit typically has 16 defects (i.e. $c_0 = 16$) given $FAR_0 = 0.0027$. In addition, assume that we are interested in detection of large shifts, hence $H = 2$ and $k = 2.085$ (see Table 4.1).

Chart Designs

The traditional charts for this example have been found in Section 4.5.1 and are displayed in Table 4.5. For the MIPL method, the calculations are as follows. From Step 1, the value of L_{max} that satisfies Equation (4.17) is equal to 8, so that $A = \{NA, 0, 1, \dots, 8\}$. For each $d = [LCL_c] \in A$, we find the corresponding f such that the $AFAR(d, f, H/16)$ of these pairs is an element of T (see Step 2). Table 4.4 shows all the possible pairs of control limits in set T (with subsets T_1 and T_2 , as defined earlier), the $AFAR$ and the percentage relative deviation from the FAR_0 for each pair (calculated in Step 3). Then using Step 4, the optimal values for the control limits of the MIPL method when the process is IC, are obtained using Table 4.4 and are given by $(d^*, f^*) = (4, 23)$ with an $AFAR$ that is only 0.01% lower than 0.0027, see the grey shading in Table 4.4.

Table 4.4: All pairs of control limits constants in set T for the synthetic c chart using the MIPL method for $c_0 = 16$, $\tau = 0.037069$ and $FAR_0 = 0.0027$

Set T_1					Set T_2				
$d = LCL_c$	$f = UCL_c - 1$	θ	$AFAR$	$ D $	$d = LCL_c$	$f = UCL_c - 1$	θ	$AFAR$	$ D $
NA	23	0.03669	0.00264	2.14%	NA	22	0.05824	0.00659	143.94%
0	23	0.03669	0.00264	2.14%	0	22	0.05824	0.00659	143.94%
1	23	0.03669	0.00264	2.13%	1	22	0.05824	0.00659	143.96%
2	23	0.03670	0.00264	2.05%	2	22	0.05826	0.00659	144.08%
3	23	0.03678	0.00266	1.64%	3	22	0.05833	0.00661	144.71%
4	24	0.02272	0.00102	62.21%	4	23	0.03709	0.00270	0.01%
5	24	0.02370	0.00111	58.89%	5	23	0.03807	0.00284	5.31%
6	24	0.02632	0.00137	49.36%	6	23	0.04069	0.00324	20.16%
7	24	0.03232	0.00205	23.90%	7	23	0.04669	0.00426	57.68%
8	25	0.03511	0.00242	10.31%	8	24	0.04430	0.00384	42.17%

Table 4.5 summarizes the three charting methods (control limits) along with the $AFAR$ and the percentage relative deviation from the FAR_0 for each pair, $AARL_0$ and $ASDRL_0$ values, respectively.

Table 4.5: Comparison among the three methods of the synthetic c chart with $c_0 = 16$

Method	(d, f)	$AFAR$	$ D $	$AARL_0$	$ASDRL_0$
k -sigma limits	(7,24)	0.00205	23.90%	486.66	514.80
Conventional probability limits	(7,25)	0.00106	60.87%	946.47	986.87
Modified improved probability limits	(4,23)	0.00270	0.01%	370.40	394.59

It is clear that, for this example, the MIPL method results in control limits with much improved IC run-length characteristics compared to the traditional 3-SL and CPL methods. For this example, the control limits $(d, f) = (4, 23)$ are the only values that ensure that we get as close as possible to the nominal ARL and $SDRL$ values.

4.6.2 Empirical comparison of the synthetic c chart methods

Similarly, as in the case of the synthetic p chart, we notice that as the $AFAR$ values converges to the FAR_0 , the $ASDRL_0$ values converges to approximately 457, 413 and 394 when $H = 47, 7$ and 2 , respectively. In addition, when the parameter c_0 is very small, all the above methods are adversely affected, because the normal approximation to the Poisson distribution is violated. From Figure 4.7 to 4.9, it can be seen that although when c_0 is small, the MIPL method

has high fluctuations in relation to the nominal values, but, as c_0 increases it has much better IC performance (in terms of $AFAR$ and $ASDRL_0$) compared to the traditional methods. Therefore, the MIPL method would be a preferred method to design a synthetic c chart than the traditional methods. Thus, it is evident from Figures 4.7 to 4.9 that Equation (4.20) ensures that the MIPL method is guaranteed to either have the same or a better IC performance when compared to the two traditional methods.

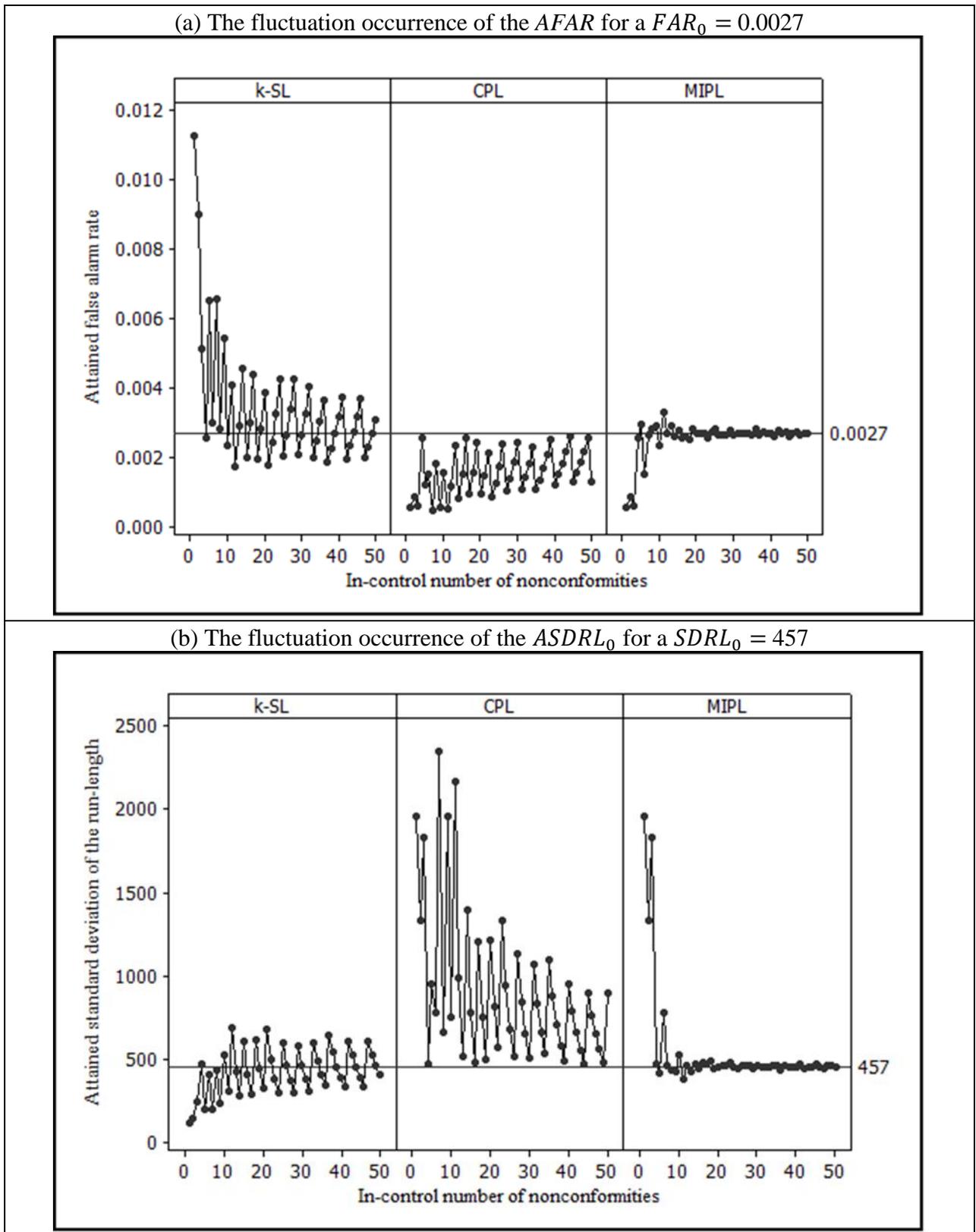


Figure 4.7: Comparison of the run-length characteristics among three methods of the synthetic c chart when the process is IC for $(H = 47, k = 2.639)$

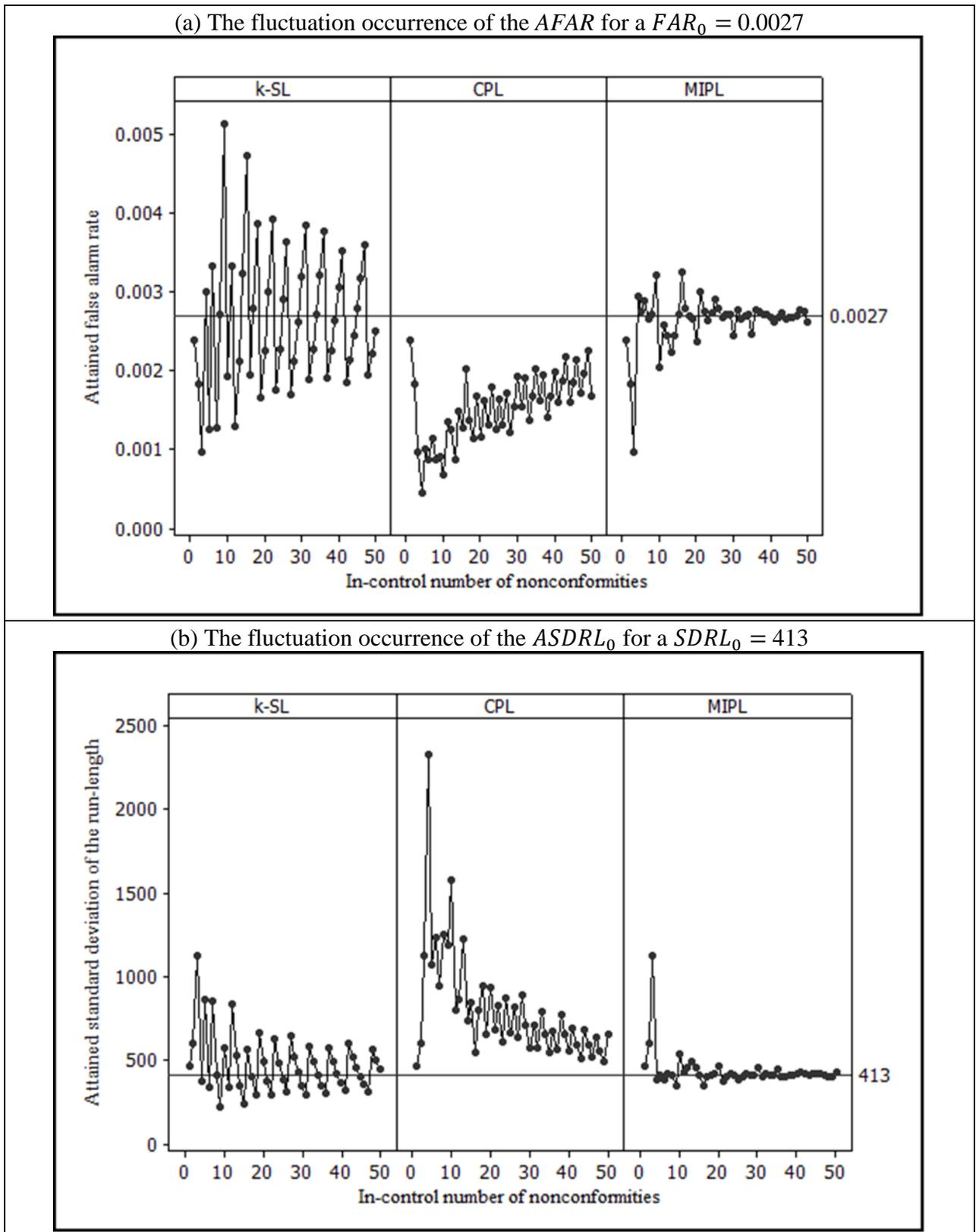


Figure 4.8: Comparison of the run-length characteristics among three methods of the synthetic c chart when the process is IC for ($H = 7, k = 2.322$)

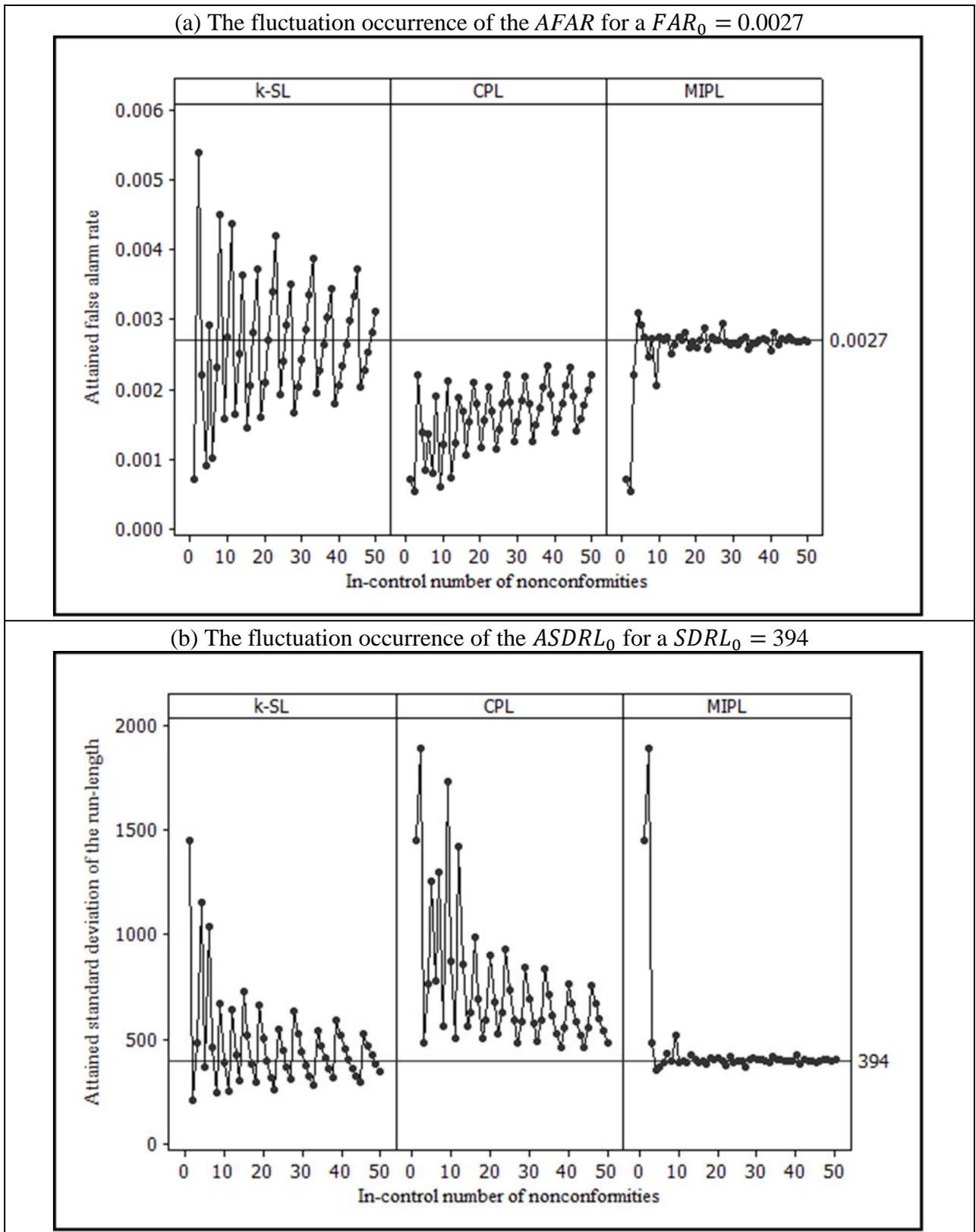


Figure 4.9: Comparison of the run-length characteristics among three methods of the synthetic c chart when the process is IC for $(H = 2, k = 2.085)$

4.6.3 Performance

Unlike Section 4.6.2, if the objective is to construct the synthetic c chart such that it has nearly ARL -unbiased control limits, we need to proceed as follows.

Step 1: Let c^* be the value of the number of nonconformities corresponding to the peak of the ARL curve, so that $ARL(d, f_i, H|c^*)$ is the value of the peak of the curve, with $(d, f_i) \in S$ defined in Equation (23). In addition, $ARL(d, f_i, H|c_0)$ is the $AARL_0$ when $c = c_0$, for $i = 1, 2$. Then construct the ARL curve for each $(d, f_i) \in S$ for some given increment shift of size δ .

Step 2: For each pair (d, f_i) in Step 1, we compute

$$q = ARL(d, f_i, H|c^*) - ARL(d, f_i, H|c_0) \quad (4.21)$$

Note that $q = 0$ if the c chart has ARL -unbiased control limit constants.

Step 3: Choose the pair $(d^\#, f^\#) \in S$ such that

$$ARL(d^\#, f^\#, H|c) = \min_{(d, f_i) \in S} ARL(d, f_i, H|c) \quad (4.22)$$

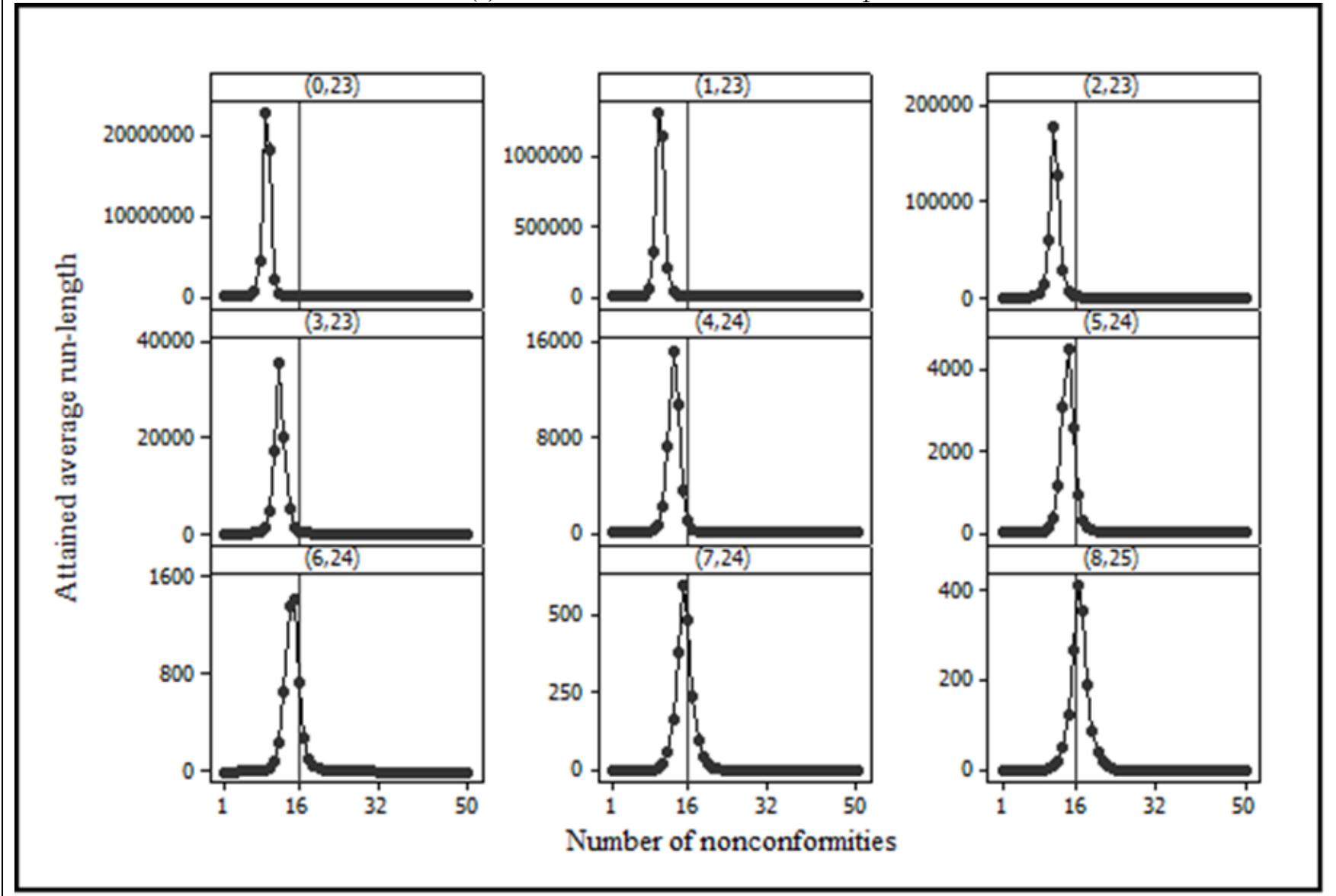
i.e. we choose the pair $(d^\#, f^\#)$ that result in the smallest value of q . Thus the nearly ARL -unbiased MIPL for the c chart are given by $LCL_c = d^\#$ and $UCL_c = f^\# + 1$. Note that, if there is more than one pair that satisfies Equation (4.22), then we must chose the pair that results in $AARL_0$ closer to ARL_0 .

To illustrate this, we use Example 4.2 to construct a synthetic c chart that will result in nearly ARL -unbiased control limits. Taking the pairs (d, f_i) in T that are given in Table 4.5 as the control limit constants, we construct the ARL curves (see Figures 4.10 (a) and (b)). It is evident that most of the control limits in set T have undesirable OOC values in addition to having a poor IC performance. By using the criteria in Equation (4.22), the pairs (8, 25) and (8, 24) result in $q = 0$, with an $AARL_0$ equal to 412.95 and 260.52, respectively. The pair (8, 25) has an $AARL_0$ that is much closer to ARL_0 .

In Figure 4.11, we plot all three competing ARL curves for the synthetic c chart and we deduce the k -SL method is ARL -biased. Furthermore, the CPL method has ARL -unbiased control limits but the $AARL_0 = 946.47$ is very large (relative to 370.4). For this example, the MIPL method results in a better OOC design compared to the two traditional methods. We observed a similar behavior for other examples that were considered.

When the normal approximation to the Poisson distribution is not satisfied, the performance of the synthetic c chart is severely degraded. Once L_{max} is a positive integer, the MIPL method is more likely to yield better OOC performance because set T has more options for the optimal pair $(d^\#, f^\#)$ compared to the k -SL and CPL methods that have only one option for (d, f) .

(a) ARL curves for control limits in set T_1



(b) ARL curves for control limits in set T_2

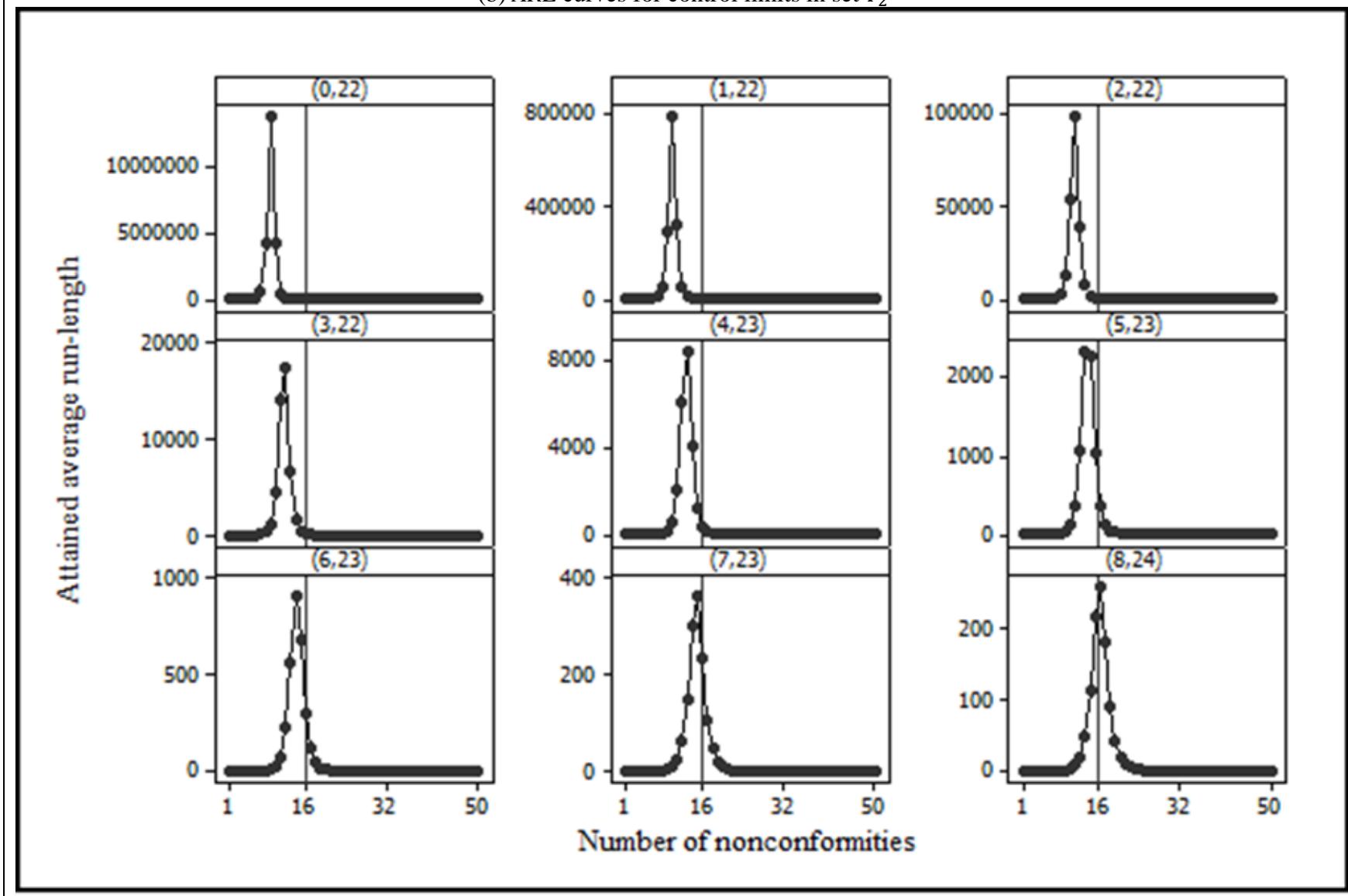


Figure 4.10: ARL curves of the control limits generated by the MIPL method for $c_0 = 16$ with $(H = 2, k = 2.085)$

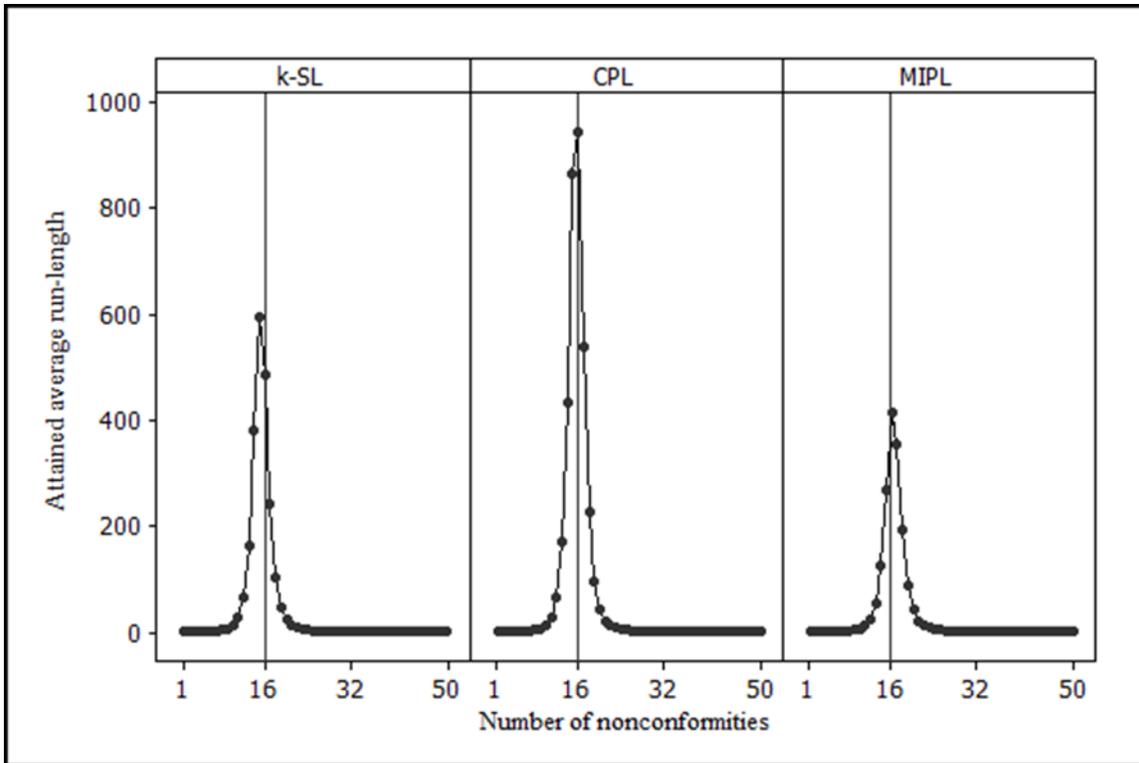


Figure 4.11: ARL curves for the three methods of the synthetic c chart when $c_0 = 16$

4.7 Concluding Remarks

In this chapter, we proposed an MIPL method for the synthetic p and c charts (and remarked on the synthetic np and u charts). The aim of this chapter was to investigate whether the MIPL method has a better IC and OOC performance compared to the k -SL and CPL methods. From this comparison, we draw the following conclusions. Firstly, when the parameters (n, p_0) and c_0 for the respective charts are small, all three methods do not have good IC and OOC performance. However, as the parameters increase, the IC run-length properties (e.g. $AFAR$ and $AARL_0$) of the MIPL method fluctuate more or less around the nominal values. Whereas, the k -SL and CPL methods have poor IC run-length characteristics even in cases when the parameters are large. Secondly, the MIPL method is more likely to have better OOC performance than the k -SL and CPL method because it creates a set of control limit constants that a practitioner can use to choose the best possible pair of control limits to design a synthetic p or c chart. Lastly, the MIPL method is time-consuming to implement compared to the k -SL and CPL methods, however, the improvement in IC and OOC performance of the control chart makes it worthwhile.

4.8 Appendix 4A: Microsoft® Excel calculations

We illustrate how the results in Example 4.1 were calculated.

3-sigma limits

The 3-sigma limits calculations in Table 2.2 were calculated as follows. The formula sheet is given by,

	A	B	C	D	E	F	G	H
1				<i>a</i>	=INT(\$B\$3*E8)			
2	<i>p</i>	0.2		<i>b</i>	=IF(B3*E9=INT(B3*E9),MIN(INT(\$B\$3*E9)-1,B3),MIN(INT(\$B\$3*E9),B3))			
3	<i>n</i>	100						
4					Check UCL Integer or NOT		<i>1-Theta</i>	<i>Theta</i>
5					=B3*E9		=IF(E1<0,BINOMDIST(E2,B3,\$B\$2,TRUE),BINOMDIST(E2,B3,\$B\$2,TRUE)-BINOMDIST(\$E\$1,B3,\$B\$2,TRUE))	=1-G4
6								
7								
8				LCL	=\$B\$2-B15*\$SQRT(\$B\$2*(1-\$B\$2)/\$B\$3)			
9				UCL	=\$B\$2+B15*\$SQRT(\$B\$2*(1-\$B\$2)/\$B\$3)			
10							<i>H</i>	
11							2	
12								
13				D	=H4*(1-((1-H4)^G11))	=1/F13		=SQRT(G18+(G21/G24))
14					=ABS((F13-0.0027)/0.0027)	<i>AFAR</i>	<i>ARL</i>	<i>SDRL</i>
15	<i>k</i>	2.085						
16					For Summation in the SDRL formula--- used to calculate the b coefficient		For Summation in the SDRL formula	
17				SUM:	=SUM(E20:E21)		<i>a</i>	
18							=(2-H4)/((1-(1-H4)^(G11))*(H4^2))	
19				<i>H=2</i>			<i>b</i>	
20			1		=D20*(1-\$H\$4)^(D20-1)			
21			2		=D21*(1-\$H\$4)^(D21-1)		=(1/(H4^2))-2*(E17)	
22								
23							<i>c</i>	
24							=(1-(1-H4)^(G11))^2	

and the corresponding value sheet is given by

	A	B	C	D	E	F	G	H
1				a	11			
2	p	0.2		b	28.0000000			
3	n	100					1-Theta	Theta
4					Check UCL Integer or NOT		0.96740	0.03260
5					28.34000			
6								
7								
8				LCL	0.11660			
9				UCL	0.28340			
10							H	
11							2	
12								
13					D	0.00209	478.41	506.29
14					22.58%	AFAR	ARL	SDRL
15	k	2.085						
16				<i>For Summation in the SDRL formula--- used to calculate the b coefficient</i>			<i>For Summation in the SDRL formula</i>	
17				SUM:	2.934809834		a	
18							28876.44481	
19				H=2				
20				1	1		b	
21				2	1.934809834		935.3604979	
22								
23							c	
24							0.004112365	

Conventional probability limits

The conventional probability limits calculations in Table 2.2 were calculated as follows. The formula sheet is given by,

	A	B	C	D	E	F	G	H
1							k	2.085
2	n	100		<i>Nom. Tail Area</i>	<i>alpha/2</i>			
3	p	0.2		0.037069	=D3/2		<i>alpha</i>	=2*(1-NORM.S.DIST(H1,TRUE))
4								
5								
6	a=LCL	<i>Left Tail</i>		b=UCL-1	<i>Right Tail</i>		a	b
7	0	=BINOMDIST(A7,\$B\$2,\$B\$3,TRUE)		0	=1-BINOMDIST(D7,\$B\$2,\$B\$3,TRUE)	11		29
8	1	=BINOMDIST(A8,\$B\$2,\$B\$3,TRUE)		1	=1-BINOMDIST(D8,\$B\$2,\$B\$3,TRUE)			
9	2	=BINOMDIST(A9,\$B\$2,\$B\$3,TRUE)		2	=1-BINOMDIST(D9,\$B\$2,\$B\$3,TRUE)			
10	3	=BINOMDIST(A10,\$B\$2,\$B\$3,TRUE)		3	=1-BINOMDIST(D10,\$B\$2,\$B\$3,TRUE)			
11	4	=BINOMDIST(A11,\$B\$2,\$B\$3,TRUE)		4	=1-BINOMDIST(D11,\$B\$2,\$B\$3,TRUE)		1-Theta	Theta
12	5	=BINOMDIST(A12,\$B\$2,\$B\$3,TRUE)		5	=1-BINOMDIST(D12,\$B\$2,\$B\$3,TRUE)		=IF(G7<0,BINOMDIST(H7,\$B\$2,\$B\$3,TRUE),BINOMDIST(H7,\$B\$2,\$B\$3,TRUE)-BINOMDIST(G7,\$B\$2,\$B\$3,TRUE))	=1-G12
13	6	=BINOMDIST(A13,\$B\$2,\$B\$3,TRUE)		6	=1-BINOMDIST(D13,\$B\$2,\$B\$3,TRUE)			
14	7	=BINOMDIST(A14,\$B\$2,\$B\$3,TRUE)		7	=1-BINOMDIST(D14,\$B\$2,\$B\$3,TRUE)		H	
15	8	=BINOMDIST(A15,\$B\$2,\$B\$3,TRUE)		8	=1-BINOMDIST(D15,\$B\$2,\$B\$3,TRUE)	2		
16	9	=BINOMDIST(A16,\$B\$2,\$B\$3,TRUE)		9	=1-BINOMDIST(D16,\$B\$2,\$B\$3,TRUE)			
17	10	=BINOMDIST(A17,\$B\$2,\$B\$3,TRUE)		10	=1-BINOMDIST(D17,\$B\$2,\$B\$3,TRUE)			
18	11	=BINOMDIST(A18,\$B\$2,\$B\$3,TRUE)		11	=1-BINOMDIST(D18,\$B\$2,\$B\$3,TRUE)			
19	12	=BINOMDIST(A19,\$B\$2,\$B\$3,TRUE)		12	=1-BINOMDIST(D19,\$B\$2,\$B\$3,TRUE)			
20	13	=BINOMDIST(A20,\$B\$2,\$B\$3,TRUE)		13	=1-BINOMDIST(D20,\$B\$2,\$B\$3,TRUE)			
21	14	=BINOMDIST(A21,\$B\$2,\$B\$3,TRUE)		14	=1-BINOMDIST(D21,\$B\$2,\$B\$3,TRUE)			
22	15	=BINOMDIST(A22,\$B\$2,\$B\$3,TRUE)		15	=1-BINOMDIST(D22,\$B\$2,\$B\$3,TRUE)			
23	16	=BINOMDIST(A23,\$B\$2,\$B\$3,TRUE)		16	=1-BINOMDIST(D23,\$B\$2,\$B\$3,TRUE)			
24	17	=BINOMDIST(A24,\$B\$2,\$B\$3,TRUE)		17	=1-BINOMDIST(D24,\$B\$2,\$B\$3,TRUE)			
25	18	=BINOMDIST(A25,\$B\$2,\$B\$3,TRUE)		18	=1-BINOMDIST(D25,\$B\$2,\$B\$3,TRUE)			
26	19	=BINOMDIST(A26,\$B\$2,\$B\$3,TRUE)		19	=1-BINOMDIST(D26,\$B\$2,\$B\$3,TRUE)			
27	20	=BINOMDIST(A27,\$B\$2,\$B\$3,TRUE)		20	=1-BINOMDIST(D27,\$B\$2,\$B\$3,TRUE)			
28	21	=BINOMDIST(A28,\$B\$2,\$B\$3,TRUE)		21	=1-BINOMDIST(D28,\$B\$2,\$B\$3,TRUE)			
29	22	=BINOMDIST(A29,\$B\$2,\$B\$3,TRUE)		22	=1-BINOMDIST(D29,\$B\$2,\$B\$3,TRUE)			
30	23	=BINOMDIST(A30,\$B\$2,\$B\$3,TRUE)		23	=1-BINOMDIST(D30,\$B\$2,\$B\$3,TRUE)			
31	24	=BINOMDIST(A31,\$B\$2,\$B\$3,TRUE)		24	=1-BINOMDIST(D31,\$B\$2,\$B\$3,TRUE)			
32	25	=BINOMDIST(A32,\$B\$2,\$B\$3,TRUE)		25	=1-BINOMDIST(D32,\$B\$2,\$B\$3,TRUE)			
33	26	=BINOMDIST(A33,\$B\$2,\$B\$3,TRUE)		26	=1-BINOMDIST(D33,\$B\$2,\$B\$3,TRUE)			
34	27	=BINOMDIST(A34,\$B\$2,\$B\$3,TRUE)		27	=1-BINOMDIST(D34,\$B\$2,\$B\$3,TRUE)			
35	28	=BINOMDIST(A35,\$B\$2,\$B\$3,TRUE)		28	=1-BINOMDIST(D35,\$B\$2,\$B\$3,TRUE)			
36	29	=BINOMDIST(A36,\$B\$2,\$B\$3,TRUE)		29	=1-BINOMDIST(D36,\$B\$2,\$B\$3,TRUE)			
37	30	=BINOMDIST(A37,\$B\$2,\$B\$3,TRUE)		30	=1-BINOMDIST(D37,\$B\$2,\$B\$3,TRUE)			
38	31	=BINOMDIST(A38,\$B\$2,\$B\$3,TRUE)		31	=1-BINOMDIST(D38,\$B\$2,\$B\$3,TRUE)			
39	32	=BINOMDIST(A39,\$B\$2,\$B\$3,TRUE)		32	=1-BINOMDIST(D39,\$B\$2,\$B\$3,TRUE)			
40	33	=BINOMDIST(A40,\$B\$2,\$B\$3,TRUE)		33	=1-BINOMDIST(D40,\$B\$2,\$B\$3,TRUE)			
41	34	=BINOMDIST(A41,\$B\$2,\$B\$3,TRUE)		34	=1-BINOMDIST(D41,\$B\$2,\$B\$3,TRUE)			
42	35	=BINOMDIST(A42,\$B\$2,\$B\$3,TRUE)		35	=1-BINOMDIST(D42,\$B\$2,\$B\$3,TRUE)			
43	36	=BINOMDIST(A43,\$B\$2,\$B\$3,TRUE)		36	=1-BINOMDIST(D43,\$B\$2,\$B\$3,TRUE)			

and the corresponding value sheet is given by

Modified improved probability limits

The following formula and value sheets show how to calculate the value of L_{max} for the MIPL method in Example 4.1.

	A	B	C	D
1				
2	<i>n</i>	100		<i>Nom. Tail Area</i>
3	<i>p</i>	0.2		=2*(1-NORM.S.DIST(2.085,TRUE))
4				
5				
6	<i>a=LCL</i>	<i>Left Tail</i>		
7	0	=BINOMDIST(A7,\$B\$2,\$B\$3,TRUE)		
8	1	=BINOMDIST(A8,\$B\$2,\$B\$3,TRUE)		
9	2	=BINOMDIST(A9,\$B\$2,\$B\$3,TRUE)		
10	3	=BINOMDIST(A10,\$B\$2,\$B\$3,TRUE)		
11	4	=BINOMDIST(A11,\$B\$2,\$B\$3,TRUE)		
12	5	=BINOMDIST(A12,\$B\$2,\$B\$3,TRUE)		
13	6	=BINOMDIST(A13,\$B\$2,\$B\$3,TRUE)		
14	7	=BINOMDIST(A14,\$B\$2,\$B\$3,TRUE)		
15	8	=BINOMDIST(A15,\$B\$2,\$B\$3,TRUE)		
16	9	=BINOMDIST(A16,\$B\$2,\$B\$3,TRUE)		
17	10	=BINOMDIST(A17,\$B\$2,\$B\$3,TRUE)		
18	11	=BINOMDIST(A18,\$B\$2,\$B\$3,TRUE)		
19	12	=BINOMDIST(A19,\$B\$2,\$B\$3,TRUE)		
20	13	=BINOMDIST(A20,\$B\$2,\$B\$3,TRUE)		
21	14	=BINOMDIST(A21,\$B\$2,\$B\$3,TRUE)		
22	15	=BINOMDIST(A22,\$B\$2,\$B\$3,TRUE)		
23	16	=BINOMDIST(A23,\$B\$2,\$B\$3,TRUE)		
24	17	=BINOMDIST(A24,\$B\$2,\$B\$3,TRUE)		
25	18	=BINOMDIST(A25,\$B\$2,\$B\$3,TRUE)		
26	19	=BINOMDIST(A26,\$B\$2,\$B\$3,TRUE)		
27	20	=BINOMDIST(A27,\$B\$2,\$B\$3,TRUE)		
28	21	=BINOMDIST(A28,\$B\$2,\$B\$3,TRUE)		
29	22	=BINOMDIST(A29,\$B\$2,\$B\$3,TRUE)		
30	23	=BINOMDIST(A30,\$B\$2,\$B\$3,TRUE)		
31	24	=BINOMDIST(A31,\$B\$2,\$B\$3,TRUE)		
32	25	=BINOMDIST(A32,\$B\$2,\$B\$3,TRUE)		
33	26	=BINOMDIST(A33,\$B\$2,\$B\$3,TRUE)		
34	27	=BINOMDIST(A34,\$B\$2,\$B\$3,TRUE)		
35	28	=BINOMDIST(A35,\$B\$2,\$B\$3,TRUE)		
36	29	=BINOMDIST(A36,\$B\$2,\$B\$3,TRUE)		
37	30	=BINOMDIST(A37,\$B\$2,\$B\$3,TRUE)		
38	31	=BINOMDIST(A38,\$B\$2,\$B\$3,TRUE)		

	A	B	C	D
1				
2	<i>n</i>	100		<i>Nom. Tail Area</i>
3	<i>p</i>	0.2		0.037069
4				
5				
6	<i>a=LCL</i>	<i>Left Tail</i>		
7	0	0.000000		
8	1	0.000000		
9	2	0.000000		
10	3	0.000001		
11	4	0.000004		
12	5	0.000019		
13	6	0.000078		
14	7	0.000277		
15	8	0.000855		
16	9	0.002334		
17	10	0.005696		
18	11	0.012575		
19	12	0.025329		
20	13	0.046912		
21	14	0.080444		
22	15	0.128506		
23	16	0.192338		
24	17	0.271189		
25	18	0.362087		
26	19	0.460161		
27	20	0.559462		
28	21	0.654033		
29	22	0.738933		
30	23	0.810913		
31	24	0.868647		
32	25	0.912525		
33	26	0.944167		
34	27	0.965848		
35	28	0.979980		
36	29	0.988751		
37	30	0.993941		
38	31	0.996870		

The rest of the steps follow as discussed in Example 4.1.

Similar calculations were done for the synthetic *c* chart in Example 4.2.

Chapter 5

Summary and Recommendations for future research

In this final chapter, we give a brief summary of the research conducted in this essay and offer concluding remarks concerning unanswered questions and future research ideas.

In this essay, we focused on statistical process control and monitoring, which is an application of a collection of statistical techniques which allows high quality products to be produced. More specifically, we focused on Shewhart-type attributes control charts to monitor count data, since, in some cases; it is not possible to quantify a quality characteristic numerically, that is, we can only classify it as either conforming or nonconforming. Moreover, we reviewed and discussed some recent developments in the area of synthetic control charts for univariate and multivariate data.

Modified improved probability limits

Our objective was to develop a more efficient method to construct attributes control charts. We illustrated that designing classical p and c charts, as well as synthetic p and c charts using the k -sigma limits (k -SL) and conventional (equal-tailed) probability limits (CPL) methods result in control limits with attained false alarm rate values that are significantly different from the target nominal value. Moreover, the control limits based on the k -SL and CPL methods are either ARL -biased, or are ARL -unbiased, however the IC ARL is very large compared to the nominal value. Thus, we offered a solution to this problem by implementing a new method of chart design called the modified improved probability limits (MIPL). The MIPL method is an adaptation and a modification of the improved probability method by Zhang et al. (2004) for a geometric chart. It was shown that the MIPL method yields control limits that result in $AFAR$ and attained $SDRL$ values that are close to the nominal values compared to the k -SL and CPL methods. Moreover, the MIPL method can be formulated such that it yields similar or better nearly- ARL unbiased control limits than the k -SL and CPL methods. We only considered the case where parameters are known (Case K) and much more remains to be done. We list a few ideas to pursue in the future.

- i. The effect of parameter estimation for both the classical and synthetic p and c charts using the MIPL method needs to be investigated. This problem has been address by Braun (1999), Chakraborti and Human (2006, 2008), Castagliola and Wu (2012) and Castagliola et al. (2013) for the k -sigma limits method.
- ii. Borror et al. (1998) showed that the EWMA c chart performs better than the c chart for small shifts and the corresponding effect of parameter estimation was investigated by Testik et al. (2006). An investigation towards the MIPL method for EWMA or CUSUM p , np , c and u charts would be interesting. In addition, the formulation of the synthetic version of these charts for both Case K and Case U could be investigated.
- iii. Attributes charts have an asymmetric run-length distribution, also in some cases there may be problems with the existence of the mean of the run-length distribution for some charts and the ARL isn't a robust measure; see, for example, Chakraborti et al. (2004) and Graham et al. (2012), hence using the median run-length to assess the performance of the chart rather than an ARL would be an interesting topic to investigate.

Synthetic control charts

Wu and Spedding (2000a) originally defined a synthetic chart as the integration of the operation of a Shewhart chart and a conforming run-length (CRL) chart. Following this, Scariano and Calzada (2009) proposed a more general approach, i.e. a synthetic chart is defined as the integration of some control charting procedure and a CRL chart. Khoo (2013) did a literature review for the univariate parametric variables synthetic charts to monitor the mean and those to monitor the variation. In this essay, we provided a more comprehensive review of the synthetic charts, by considering variables (parametric and nonparametric) and attributes synthetic charts for both univariate and multivariate data. A number of important topics for synthetic charts have already been investigated, however, more work still remains to be done. In addition to the suggestions for future research given in Khoo (2013), here is a summary of some topics/questions about synthetic charts that have not yet been addressed or have only been partially answered.

- i. Zhang et al. (2011) gave a thorough account of the run-length properties for a synthetic \bar{X} chart for Case U under the zero-state mode, however, run-length properties for the synthetic T^2 chart for Case U has not yet been explored in detail. The reader can start by reading Champ et al. (2005) as they provided properties of the T^2 chart for Case U.
- ii. The effect of parameter estimation for the synthetic \bar{X} chart using the steady-state analysis needs to be investigated. In addition, other than the synthetic \bar{X} chart (in Zhang et al. (2011)) and the synthetic p , np , c and u charts (in Castagliola et al. (2013)), the effect of parameter estimation has not been investigated for other synthetic-type charts.
- iii. Huang and Chen (2005) suggested investigating the effect of autocorrelation in monitoring process dispersion, since many authors typically use the assumption of independence between the monitored quality characteristics. Note that, Machado et al. (2009) has investigated this problem for bivariate data, however, more work still remains to be done.
- iv. Huang and Chen (2005) suggested investigation of synthetic charts based on the moving range for individual observations. While on the subject of individual observations, one other synthetic chart that may be investigated is one based on the moving average for both univariate and multivariate processes. See Ghute and Shirke (2013) for a multivariate Hotelling's T^2 chart based on the moving average. Combining the operation of this chart and a CRL chart will yield a multivariate synthetic moving average chart.
- v. Most of the multivariate synthetic charts are based on the assumption of normality. Thus, robustness of these multivariate charts to non-normality still has to be studied for both zero-state and steady-state modes.
- vi. Generalized synthetic charts (GSC) for Case K under zero-state mode has been derived in Scariano and Calzada (2009) and has been used by a number of authors in

this essay. However, GSC for Case U under both zero-state and steady-state modes needs to be formulated.

- vii. The individual or combined VSS and VSI schemes are generally known to be more statistically efficient than FSS and FSI schemes (see Costa (1997)). Synthetic charts based on the combined scheme have not been investigated and cases when VSS and/or VSI is applied in parameter estimation.
- viii. Economic and economic-statistical designs of the synthetic-type charts (other than those with \bar{X} and T^2 as sub-charts) have not yet been investigated. For example, the starting point to study economic design of the multivariate synthetic np chart is given in Jolayemi (2000) where the economic design of the standard multivariate np chart is studied.
- ix. Khoo et al. (2011) showed that the double sampling synthetic \bar{X} chart has a significantly better performance than its FSS counterpart. However, this chart was proposed for the zero-state mode; hence it would be interesting to study the performance of this chart under the steady-state mode.
- x. Yen et al. (2013) considered synthetic charts for time-between events when parameters are known; the effect of parameter estimation for this scenario still needs to be investigated.
- xi. More nonparametric (NP) synthetic charts need to be formulated. As stated earlier, only synthetic charts based on the sign and signed-rank statistics have been proposed to date. One could study, for example, a synthetic chart based on a Mann-Whitney statistic, Mood statistic, precedence or exceedence statistics, Conover's squared rank test for variance, etc. Furthermore, a NP synthetic chart to jointly monitor the location and dispersion can be formulated by integrating the operation of the chart proposed in Mukherjee and Chakraborti (2012) or Chowdhury et al. (2013) with the CRL chart. Similarly, multivariate NP synthetic charts can be investigated by using the sub-charts already proposed in Das (2009), Zou and Tsung (2011) and Li et al. (2013).

- xii. NP Shewhart-type charts are generally known to be efficient for the detection of large shifts. NP schemes with CUSUM or EWMA procedures, to monitor small shifts, were proposed in Bakir and Reynolds (1979), Amin and Searcy (1991), Amin et al. (1995), Li et al. (2010) and further studied in Graham et al. (2011a, b) for the sign, signed-rank and Wilcoxon rank-sum tests. These schemes may be integrated with the CRL chart to develop more efficient NP synthetic charts.
- xiii. Until now, only a few synthetic-type charts have been proposed to jointly monitor the mean and variance (for parametric and nonparametric charts). Reviews by Cheng and Thaga (2006) and McCracken and Chakraborti (2013) give a number of charting procedures for this purpose whose operation can be integrated with the CRL chart to investigate whether this will yield charts that can efficiently detect process disturbances better than the existing methods.

Research outputs

Next we list the research outputs associated with this dissertation. This includes local conferences where papers have been presented, departmental seminars and papers in progress.

National conference (presentations)

- i. Shongwe, S.C., Chakraborti, S. and Graham, M.A. (2012). “Improved probability limits design for attributes data.” The 54th annual conference of the South African Statistical Association (SASA), Nelson Mandela Metropolitan University (NMMU), Port Elizabeth, 5 – 9 November 2012.
- ii. Shongwe, S.C., Chakraborti, S., Graham, M.A. (2013). “Modified improved probability limits for the synthetic c chart.” The 55th annual conference of the South African Statistical Association (SASA), University of Limpopo (UL), Polokwane, 4 – 8 November 2013.

Departmental seminars (presentations)

- i. Shongwe, S.C., Chakraborti, S. and Graham, M.A. (2012). “Improved probability limits design for attribute data.” Departmental Seminar, University of Pretoria, Pretoria, South Africa, Nov 2012.
- ii. Shongwe, S.C., Chakraborti, S. and Graham, M.A. (2013). “Modified improved probability limits design for synthetic c control charts.” Departmental Seminar, University of Pretoria, Pretoria, South Africa, Nov 2013.

Papers in progress

- i. Shongwe, S.C., Chakraborti, S. and Graham, M.A. (2014). “Modified improved probability limits for the p and c control charts.”
- ii. Shongwe, S.C., Chakraborti, S. and Graham, M.A. (2014). “Modified improved probability limits for the synthetic p and c control charts.”
- iii. Shongwe, S.C., Chakraborti, S. and Graham, M.A. (2014). “Synthetic control charts: An overview.”
- iv. Shongwe, S.C., Chakraborti, S. and Graham, M.A. (2014). “Comparison of four variables control charts to monitor the process mean.”

References

- Abbasi, S.A., Riaz, M. and Miller, A. (2012). “Enhancing the performance of CUSUM scale chart.” *Computers & Industrial Engineering*, **63** (2), 400-409.
- Acosta-Mejia, C.A. (1999). “Improved p charts to monitor process quality”. *IIE Transactions*, **31** (6), 509-516.
- Aebtarm, S. and Bouguila, N. (2011). “An empirical evaluation of attributes control charts for monitoring defects.” *Expert Systems with Applications*, **38** (6), 7869-7880.
- Amin, R. W. and Searcy, A.J. (1991). “A nonparametric exponentially weighted moving average control scheme”. *Communications in Statistics—Simulation and Computation*, **20** (4), 1049–1072.
- Amin, R. W., Reynolds, Jr., M. R. and Bakir, S. (1995). “Nonparametric quality control charts based on the sign statistic”. *Communications in Statistics—Theory and Methods*, **24** (6), 1597–1623.
- Aparisi, F. and García-Díaz, C. (2007). “Design and optimization of EWMA control chart for in-control indifference and out-of-control regions”. *Computers & Operations Research*, **34** (7), 2096–2108.
- Aparisi, F. and de Luna M.A. (2009a). “Synthetic control charts optimized for in-control and out-of-control regions”. *Computers and Operations research*, **36** (12), 3204-3214.
- Aparisi, F. and de Luna, M.A. (2009b). “The design and performance of the multivariate synthetic T^2 control chart”. *Communications in Statistics – Theory and Methods*, **38** (2), 173-192.
- Bakir, S. T. (2006). “Distribution-free quality control charts based on signed-rank-like statistics”. *Communications in Statistics – Theory and Methods*, **35** (4), 743-757.

- Bakir, S.T. (2011). *Distribution-Free (Nonparametric) Statistical Quality Control Charts: A Concise Summary Part I (1920's-2000)*.
- Bakir, S. T. and Reynolds, Jr., M.R. (1979). "A nonparametric procedure for process control based on within-group ranking". *Technometrics*, **21** (2), 175-183.
- Bennett, B. M. (1964). "A bivariate signed rank test". *Journal of the Royal Statistical Society, Series B*, **26**, 457-461.
- Bersimis, S., Psarakis, S. and Panaretos, J. (2007). "Multivariate statistical process control charts: An overview". *Quality and Reliability Engineering International*, **23**, 517-543.
- Borror, C.M., Champ, C.W. and Rigdon, S.E. (1998). "Poisson EWMA control charts". *Journal of Quality Technology*, **30** (4), 352-361.
- Bourke, P. D. (1991). "Detecting a shift in fraction nonconforming using run-length control charts with 100% inspection". *Journal of Quality Technology*, **23** (3), 225-238.
- Bourke, P. D. (2008). "Comparisons for the synthetic control charts for detecting increases in fraction nonconforming". *Journal of Quality Technology*, **40** (4), 461-475.
- Braun, W.J. (1999). "Run length distributions for the estimated attributes charts". *Metrika*, **50** (2), 121-129.
- Brook, D. and Evans, D.A. (1972). "An approach to the probability distribution of CUSUM run length." *Biometrika*, **59** (3), 539-549.
- Calzada, M.E. and Scariano, S.M. (2001). "The robustness of the synthetic control chart to non-normality". *Communications in Statistics - Simulation and Computations*, **30** (2), 311-326.
- Calzada, M.E. and Scariano, S.M. (2013a). "A synthetic t and synthetic EWMA t charts". *Quality Technology and Quantitative Management*, **10** (1), 37-56.

Calzada, M.E. and Scariano, S.M. (2013b). “A synthetic control chart for the coefficient of variation”. *Journal of Statistical Computation and Simulation*, **83** (5), 851-865.

Castagliola, P. and Khoo, M.B.C. (2009). “A synthetic scaled weighted variance control chart for monitoring the process mean of skewed populations”. *Communications in Statistics – Simulation and Computation*, **38** (8), 1659-1674.

Castagliola, P., Celano, G. and Psarakis, S. (2011). “Monitoring the coefficient of variation using EWMA charts”. *Journal of Quality Technology*, **43** (3), 249-265.

Castagliola, P. and Wu, S. (2012). “Design of the c and np charts when the parameters are estimated”. *International Journal of Reliability, Quality and Safety Engineering*, **19** (2), 1250010.

Castagliola, P., Wu, S., Khoo, M.B.C. and Chakraborti, S. (2013). “Synthetic Phase II Shewhart-type attributes control charts when process parameters are estimated”. *Quality and Reliability Engineering International*, DOI: 10.1002/qre.1576.

Chakraborti, S., Van der Laan, P. and Bakir, S. (2001). “Nonparametric control charts: an overview and some results”. *Journal of Quality Technology*, **33** (3), 304–315.

Chakraborti, S. and Human, S.W. (2006). “Parameter estimation and performance of the p -chart for attributes data”. *IEEE Transactions on Reliability*, **55** (3), 559-566.

Chakraborti, S. and Graham, M. A. (2007). “Nonparametric control charts”. In: *Encyclopedia of Statistics in Quality and Reliability*, Vol. 1. New York: John Wiley, 415–429.

Chakraborti, S. and Eryilmaz, S. (2007). “A nonparametric Shewhart-type signed-rank control chart based on runs”. *Communications in Statistics—Simulation and Computation*, **36** (2), 335–356.

Chakraborti, S. and Human, S.W. (2008). “Properties and performance of the c -chart for attributes data.” *Journal of Applied Statistics*, **35** (1), 89-100.

Chakraborti, S., Human, S.W. and Graham, M.A. (2009). "Phase I statistical process control charts: An overview and some results." *Quality Engineering*, **21** (1), 52-62.

Chakraborti, S. and van de Wiel, M. (2008): "A nonparametric control chart based on the Mann-Whitney statistic" In *Beyond Parametrics in Interdisciplinary Research, Festschrift to P. K. Sen, Institute of Mathematical Statistics Collections*, Vol. 1, Edited by N. Balakrishnan, E. Pena, and M. J. Silvapulle.

Chakraborti, S., Human, S.W. and Graham, M.A. (2011). "Nonparametric (distribution-free) quality control charts." In *Handbook of Methods and Applications of Statistics: Engineering, Quality Control, and Physical Sciences*. N. Balakrishnan, Ed., 298-329, John Wiley & Sons, New York.

Champ, C. and Woodall, W.H. (1987). "Exact results for Shewhart control charts with supplementary run rules". *Technometrics*, **29** (4), 393-399.

Champ, C., Jones-Farmer, L.A. and Rigdon, S.E. (2005). "Properties of the T^2 control charts when parameters are estimated". *Technometrics*, **47** (4), 437-445.

Chen, F.L. and Huang H.J. (2005). "A synthetic control chart for monitoring process dispersion with sample range". *International Journal of Advanced manufacturing Technology*, **26** (7-8), 842-851.

Chen, F.L. and Huang H.J. (2006). "Variable sampling interval synthetic control charts for jointly monitoring process mean and standard deviation". *International Journal of Industrial Engineering: Theory Applications and Practice*, **13** (2), 136-146.

Cheng, S.W. and Thaga, K. (2006). "Single variables control charts: an overview". *Quality and Reliability Engineering International*, **22** (7), 811-820.

Chong, Z.L., Khoo, M.B.C. and Castagliola, P. (2014). "Synthetic double sampling np control chart for attributes". Submitted.

Chowdhury, S., Mukherjee, A. and Chakraborti, S. (2013). “A new distribution-free control chart for joint monitoring of unknown location and scale parameters of continuous distributions”. *Quality and Reliability Engineering International*, DOI: 10.1002/qre.1488.

Chung, K.J. (1990). “A simplified procedure for the economic design of \bar{X} charts”. *International Journal of Production Research*, **28** (7), 1239-1246.

Costa, A.F.B. (1997). \bar{X} charts with variable sample sizes and sampling intervals. *Journal of Quality Technology*, **29** (2), 197–204.

Costa, A.F.B. and Rahim, M.A. (2006). “A synthetic control chart for monitoring the process mean and variance”. *Journal of Quality in Maintenance Engineering*, **12** (1), 81-88.

Costa, A.F.B. and Machado, M.A.G. (2007). “Synthetic control charts with two-stage sampling for monitoring bivariate processes”. *Pesquisa operacional*, **27** (1), 117-130.

Costa, A.F.B., de Magalhaes, M.S., Epprecht, E.K. (2008). “Monitoring the process mean and variance using a synthetic chart with two-stage testing”. *International Journal of Production Research*, **47** (18), 5067-5086.

Das, N. (2008). “A note on efficiency of nonparametric control chart for monitoring process variability”. *Economic Quality Control*, **23** (1), 85-93.

Das, N. (2008). “Nonparametric control chart for controlling variability based on rank test”. *Economic Quality Control*, **23** (2), 227–242.

Das, N. and Bhattacharya, A. (2008). “A new nonparametric control chart for controlling variability”. *Quality Technology and Quantitative Management*, **5** (4), 351-361.

Das, N. (2009). “A new multivariate nonparametric control chart based on sign chart”. *Quality Technology and Quantitative Management*, **6** (2), 155-169.

Davis, R.B. and Woodall, W.H (2002). "Evaluating and improving the synthetic control chart". *Journal of Quality Technology*, **34** (2), 200-208.

Fang, Y.Y., Khoo, M.B.C. and Lee, M.H. (2013). "Synthetic-type control charts for time-between-events monitoring". *PLoS ONE* **8** (6): e65440.doi: 10.1371/journal.pone.0065440.

Fu, J.C. and Lou, W.Y.W. (2003). *Distribution Theory of Runs and Patterns and Its Applications: A Finite Markov Chain Imbedding Approach*. Singapore: World Scientific Publishing.

Ghute, V.B. and Shirke D.T. (2007). "Joint monitoring of multivariate process using synthetic control charts". *International Journal of Statistics and Management System*, **1** (2), 129-141.

Ghute V.B. and Shirke, D.T. (2008a). "A multivariate synthetic control chart for monitoring process dispersion". *Quality Technology and Quantitative Management*, **5** (3), 271-288.

Ghute V.B. and Shirke, D.T. (2008b). "A multivariate synthetic control chart for monitoring process mean vector". *Communications in Statistics – Theory and Methods*, **37** (13), 2136-2148.

Ghute, V.B. and Shirke, D.T. (2012). "A nonparametric signed-rank control chart for bivariate process location". *Quality Technology and Quantitative Management*, **9** (4), 317-328.

Ghute, V.B. and Shirke, D.T. (2013). "A multivariate moving average control chart for mean vector". *Journal of Academia and Industrial Research*, **1** (12), 795-800.

Graham, M. A., Chakraborti, S. and Human, S. W. (2011). "A nonparametric EWMA sign chart for location based on individual measurements". *Quality Engineering*, **23** (3), 227-241.

Graham, M.A., Chakraborti, S. and Human, S. W. (2011). "A nonparametric exponentially weighted moving average signed-rank chart for monitoring location". *Computational Statistics and Data Analysis*, **55** (8), 2490–2503.

Graham, M.A. (2013). "Contributions to the theory and applications of univariate distribution-free Shewhart, CUSUM and EWMA control charts." Unpublished PhD dissertation, Department of Statistics, University of Pretoria.

Haridy, S., Wu, Z., Khoo, M.B.C. and Yu, F.-J. (2012). "A combined synthetic and np scheme for detecting increases in fraction nonconforming". *Computers and Industrial Engineering*, **62** (4), 979-988.

Haridy, S., Wu, Z., Abhary, K., Castagliola, P. and Shamsuzzaman, M. (2013). "Development of a multiattribute synthetic- np chart". *Journal of Statistical Computation and Simulation*, DOI: 10.1080/00949655.2013.769541.

Hawkins, D.M. (1987). "Self-starting CUSUM charts for location and scale." *The Statistician*, **36** (4), 299-316.

Hawkins, D.M. (1993). "Cumulative sum control charting: An underutilized SPC tool." *Quality Engineering*, **5** (3), 463-477.

Hawkins, D.M. and Olwell, D.H. (1998). *Cumulative sum charts and charting for quality improvement*. Springer-Verlag, New York.

Huang H.J. and Chen, F.L. (2005). "A synthetic control chart for monitoring process dispersion with sample standard deviation". *Computers and Industrial Engineering*, **49** (2), 221-240.

Human, S.W. and Graham, M.A. (2007). "Average run-lengths and operating characteristic curves". *Encyclopedia of Statistics in Quality and Reliability*, **1**, 159-168, New York: John Wiley.

Human, S.W. (2009). “Univariate parametric and nonparametric statistical quality control techniques with estimated process parameters.” Published PhD dissertation, Department of Statistics, University of Pretoria.

Jolayemi, J.K. (1999). “A statistical model for the design of multiattribute control charts”. *Indian Journal Statistics*, **61**, 351–365.

Jolayemi, J.K. (2000). “An optimal design of multiattribute control charts for processes subject to multiplicity of assignable causes”. *Applied Mathematics and Computation*, **114** (2-3), 187-203.

Khilare, S.K. and Shirke, D.T. (2010). “A nonparametric synthetic control chart using sign statistic”. *Communications in Statistics – Theory and Methods*, **39** (18), 3282-3293.

Khilare, S.K. and Shirke, D.T. (2012). “Nonparametric synthetic control charts for process variation”. *Quality and Reliability Engineering International*, **28** (2), 193-202.

Khoo, M.B.C., Wu, Z. and Atta A.M.A. (2008). “A synthetic control chart for monitoring the process mean of skewed populations based on the weighted variance method”. *International Journal of Reliability, Quality and Safety Engineering*, **15** (3), 217-245.

Khoo, M.B.C., Atta A.M.A. and Wu, Z. (2009). “A multivariate synthetic control chart for monitoring the process mean vector of skewed populations using weighted standard deviations”. *Communications in Statistics – Simulation and Computation*, **38** (7), 1493-1518.

Khoo, M.B.C., Lee, H.C., Wu, Z., Chen, C.H. and Castagliola, P., (2011a). “A synthetic double sampling control chart for the process mean”. *IIE Transactions*, **43** (1), 23-38.

Khoo, M.B.C., Wong, V.H., Wu, Z. and Castagliola, P. (2011b). “Optimal designs of the multivariate synthetic chart for monitoring the process mean vector based on the median run length”. *Quality and Reliability Engineering International*, **27** (8), 981-997.

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 the median run length”. *IIE Transactions*, **44** (9), 765-779.

Khoo, M.B.C., Wu, Z., Castagliola, P., and Lee, H.C. (2013). “A multivariate synthetic double sampling T^2 control chart”. *Computers and Industrial Engineering*, **64** (1), 179-189.

Khoo, M.B.C. (2014). “Recent developments on synthetic control charts”. Submitted.

Kritzinger, P. (2011). “Improved nonparametric control charts for location based on runs-rules”. Unpublished MSc dissertation, Department of Statistics, University of Pretoria.

Kusukawa, E., and Ohta, H. (2005). “A synthetic chart to monitor the defect rate for high-yield processes”. *An International Journal of Industrial Engineering and Management Systems*, **4** (2), 158-164.

Kusukawa, E., Kotani, T. and Ohta, H. (2008). “A synthetic exponentially weighted moving average chart for high-yield processes”. *An International Journal of Industrial Engineering and Management Systems*, **7** (2), 101-112.

Latouche, G. and Ramaswami, V. (1999). *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA SIAM, Philadelphia, PA.

Lee M.H. (2012). “The design of the multivariate synthetic exponentially weighted moving average control chart”. *Communications in Statistics – Simulation and Computation*, **41** (10), 1785-1793.

Lee, M.H., Khoo, M.B.C. and Xie, M. (2013). “An optimal control procedure based on multivariate synthetic cumulative sum”. *Quality and Reliability Engineering International*, **29** (4), DOI: 10.1002/qre.1533.

Lee, M.H. and Khoo, M.B.C. (2013). “The synthetic mean square error control chart”. *Communications in Statistics – Simulation and Computation*, DOI: 10.1080/03610918.2012.735321.

Lee, M.H. and Khoo, M.B.C. (2013). “Multivariate synthetic $/S/$ control chart with variable sampling interval”. *Communications in Statistics – Simulation and Computation*, Accepted.

Li, S., Tang, L.C. and Ng, S. (2010). “Nonparametric CUSUM and EWMA control charts for detecting mean shifts”. *Journal of Quality Technology*, **42** (2), 209-226.

Li, J., Zhang, X. and Jeske, D.R. (2010). “Nonparametric multivariate CUSUM control charts for location and scale changes”. *Journal of Nonparametric Statistics*, **25** (1), 1-20.

Machado, M.A.G., Costa, A.F.B. and Marins, F.A.S. (2009). “Control charts for monitoring the mean vector and covariance matrix of bivariate processes”. *International Journal of Advanced Manufacturing Technology*, **45** (7-8), 772-785.

Machado, M.A., Costa, A.F. and Rahim, M.A. (2009). “The synthetic control chart based on two sample variances for monitoring the covariance matrix”. *Quality and Reliability Engineering International*, **25** (5), 595-606.

Mishima, N., Kusakawa, E., and Ohta, H. (2002). “A synthetic chart for high-yield processes”. *Proceedings of the 7th International Pacific Conference on Manufacturing & Management, Bangkok, Thailand*, 799-804.

Montgomery, D.C. (2013). *Statistical Quality Control: A Modern Introduction*, 7th ed., John Wiley & Sons, Singapore Pte. Ltd.

Mukherjee, A. and Chakraborti, S. (2012). “A distribution-free control chart for joint monitoring of location and scale”. *Quality and Reliability Engineering International*, **28** (3), 335–352.

Page, E.S. (1954). “Continuous inspection schemes.” *Biometrika*, **41** (1), 100-115.

Pawar, V.Y. and Shirke, D.T. (2010). "A nonparametric Shewhart-type synthetic control chart". *Communications in Statistics – Simulation and Computation*, **39** (8), 1493-1505.

Pignatiello, Jr. J.J., Acosta-Mejia, C. A. and Rao, B.V. (1995). "The performance of control charts for monitoring process dispersion". *Proceedings of the 4th Industrial Engineering Research Conference*, 320-328.

Psarakis, S., Vyniou, A.K., Castagliola, P. (2013). "Some recent developments on the effects of parameter estimation on control charts". *Quality Reliability Engineering International*, DOI: 10.1002/qre.1556.

Rajmanya, S.V. and Ghute V.B. (2013a). "A synthetic control chart for monitoring variability". *Quality Reliability Engineering International*, DOI: 10.1002/qre.1551.

Rajmanya, S.V. and Ghute V.B. (2013b). "The synthetic D chart under non-normality". *International Journal of Engineering Research and Technology*, **2** (4), ISSN: 2278-0181.

Ruggeri, F., Kenett, R.S. and Faltin, F.W. (2007). "Exponentially weighted moving average (EWMA) control chart." *Encyclopedia of Statistics in Quality and Reliability*, **2**, 633-639, John Wiley, New York.

Scariano, S.M. and Calzada M.E. (2003). "A note on the lower-sided synthetic chart for exponentials". *Quality Engineering*, **15** (4), 677-680.

Scariano, S.M. and Calzada, M.E. (2009). "The generalized synthetic chart". *Sequential Analysis*, **28** (1), 54-68.

Sim, C.H. (2003). "Combined \bar{X} and CRL charts for the gamma process". *Computational Statistics*, **18** (4), 547-563.

Steiner, S.H. (1999). "Confirming sample control chart". *International Journal of Production Research*, **37** (4), 737-748.

Szarka III, J.L. and Woodall W.H. (2011). "A review and perspective on surveillance of Bernoulli processes". *Quality and Reliability Engineering International*; **27** (6), 735-752.

Testik, M.C., McCullough, B.D. and Borrer, C.M. (2006). "The effect of estimated parameters on Poisson EWMA control charts". *Quality Technology and Quantitative Management*, **3** (4), 513-527.

Woodall, W.H. (1997). "Control charts based on attribute data: bibliography and review". *Journal of Quality Technology*, **29** (2), 172-183.

Woodall, W.H. and Montgomery, D.C. (1999). "Research issues and ideas in statistical process control." *Journal of Quality Technology*, **31** (4), 376-386.

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

Wu, Z. and Spedding, T.A. (2000b). "Implementing synthetic control charts". *Journal of Quality Technology*, **32** (1), 75-78.

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.

Wu, Z and Yeo, S.H. (2001). "Implementing synthetic control charts for attributes". *Journal of Quality Technology*, **33** (1), 112-114.

Wu, Z. and Wang, Q. (2007). "An np control chart using double inspections." *Journal of Applied Statistics*, **34** (7), 843-855.

Wu, Z., Ou, Y.J., Castagliola, P. and Khoo, M.B.C. (2010). "A combined synthetic and X chart for monitoring the process mean". *International Journal of Production Research*, **48** (24), 7423-7436.

Yang, Z., Xie, M., Kuralmani, V. and Tsui, K. L. (2002). "On the performance of geometric chart with estimated control limits". *Journal of Quality Technology*, **34** (4), 448-458.

Yeong, W.C., Khoo, M.B.C., Wu, Z. and Castagliola, P. (2012). “Economically optimum design of a synthetic \bar{X} chart”. *Quality and Reliability Engineering International*, **28** (7), 725-741.

Yeong, W.C., Khoo, M.B.C., Lee, M.H. and Rahim, M.A. (2013a). “Economic and economic statistical designs of the synthetic \bar{X} chart using loss functions”. *European Journal of Operations Research*, **228** (3), 571-581.

Yeong, W.C. and Khoo M.B.C. (2013). “Sensitivity analyses of the design parameters of the economically optimal synthetic chart”. *3rd International Conference on Management, Economics and Social Sciences*, Jan. 8-9, Kuala Lumpur (Malaysia).

Yeong, W.C., Khoo, M.B.C., Lee, M.H. and Rahim, M.A. (2014a). “Economically optimal design of a multivariate synthetic T^2 chart”. *Communications in Statistics – Simulation and Computation*. 43 (6), 1333-1361.

Yeong, W.C., Khoo, M.B.C., Shamsuzzaman, M. and Castagliola, P. (2014b). “Economic and economic statistical designs of the multivariate synthetic T^2 chart using loss functions”. Submitted.

Yeong, W.C., Khoo, M.B.C., Shamsuzzaman, M. and Castagliola, P. (2014c). “Sensitivity analysis of the economic and economic statistical designs of the multivariate synthetic T^2 chart”. Submitted.

Zhang, L., Govindaraju, K., Bebbington, M. and Lai, C.D. (2004). “On the statistical design of geometric control charts.” *Quality Technology and Quantitative Management*, **1** (2), 233-243.

Zhang, L., Chen, G. and Castagliola, P. (2009). “On t and EWMA t charts for monitoring changes in the process mean”. *Quality and Reliability Engineering International*, **25** (8), 933-945.

Zhang, Y., Castagliola, P., Wu, Z. and Khoo, M.B.C. (2011). “The synthetic \bar{X} chart with estimated parameters”. *IIE Transactions*, **43** (9), 676-687.

Zou, C. and Tsung, F. (2011). “A multivariate sign EWMA control chart”. *Technometrics*, **53** (1), 84-97.