UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Theory and applications of univariate distribution-free Shewhart, CUSUM and EWMA control charts

by

Marien Graham

Submitted in partial fulfilment of the requirements for the degree

MSc (Mathematical Statistics)

in the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

April 2008

Declaration

I declare that the dissertation/thesis, which I hereby submit for the degree MSc (Mathematical 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 thank Professor C.F. Smit for his thorough and insightful comments leading to substantial improvements to this thesis.

I would also like to thank Professor S. Chakraborti and Mr. S.W. Human for all the valuable discussions and for suggesting helpful improvements.

Finally, I would like to thank my husband, Walter, for his love, support and encouragement.

Summary

Statistical quality control charts originated in the late 1920's by Shewhart (1926, 1931 and 1939). Their applications in various disciplines have been ever-increasing. Although most control charts are distribution-based, recent literature witnessed the development of a considerable number of distribution-free or nonparametric control charts.

The purpose of this thesis is to present the concepts and introduce the researcher to the essentials of univariate nonparametric control charts. Various properties of nonparametric control charts are comprehensively discussed and concepts are clearly explained. Proofs and detailed calculations have been given to help the reader to study and understand the subject more thoroughly. This text contains a wide variety of illustrative examples to give an overall picture of how nonparametric control charts are used. Both simulated and real data examples have been integrated throughout the text. Since most practical problems are too large to be solved using hand calculations, some type of statistical software package is required to solve these problems. There are several excellent statistical packages available and in this thesis we make use of Microsoft Excel, SAS, Minitab, Mathcad and Mathematica to construct (almost all) the tables in this thesis. We point out that a number of Mathematica programs are provided by Chakraborti and Van de Wiel (2003) by means of the website www.win.tue.nl/~markvdw.

The aim throughout is to convey the concepts of univariate nonparametric control charts in a way that readers will find attractive and interesting. Since the majority of nonparametric procedures, to be distribution-free, require a continuous population, only variables control charts are covered. We only consider control charts for monitoring the location of a process, since very few nonparametric charts are available for monitoring the spread. In this thesis we consider the three main classes of control charts: the Shewhart, CUSUM and EWMA control charts and their refinements. The text is divided into several chapters. An introduction to nonparametric control charts is presented in Chapter 1. A discussion of some of the advantages of nonparametric control charts is included while pointing out some of the disadvantages. In Chapter 2 we describe the Shewhart-, CUSUM- and EWMA-type sign control charts with (and without) warning limits. In Chapter 3 we describe the Shewhart-, CUSUM- and EWMA-type signed-rank control charts with (and

without) runs-type signalling rules. The Shewhart-type sign-like control chart with (and without) signalling rules is considered in Chapter 4. In Chapter 5 we consider the Shewhart-type signed-rank-like control chart. Finally, in Chapter 6 we consider the Shewhart- and CUSUM-type Mann-Whitney-Wilcoxon control charts. We considered decision problems under both Phase I and Phase II (see Section 1.5 for a distinction between the two phases). In all the sections of this thesis we considered Phase II process monitoring, except in Section 6.2 where a CUSUM-type control chart for the preliminary Phase I analysis of individual observations based on the Mann-Whitney two-sample test is proposed. In the last chapter we have some concluding remarks along with some ideas for future research.

Table of contents

1. Introduction **11**

1.1. Notation 11

1.2. Distribution of chance causes 12

1.3. Nonparametric or distribution-free 12

1.4. Nonparametric control charts 13

1.5. Terminology and formulation 14

1.6. Shewhart-type charts 15

1.7. CUSUM-type charts 16

1.8. EWMA-type charts 16

Section A: Monitoring the location of a process when the target location is specified (Case K)

2. Sign control charts **17**

2.1. The Shewhart-type control chart 17

2.1.1. Introduction 17
2.1.2. Definition of the sign test statistic 17
2.1.3. Plotting statistic 19
2.1.4. Determination of control limits 21
2.1.5. Run length distribution 27

2.1.6. One-sided control charts 29
2.1.6.1.Lower one-sided control charts 29
2.1.6.2.Upper one-sided control charts 31
2.1.7. Two-sided control charts 37
2.1.8. Summary 38

2.2. The Shewhart-type control chart with warning limits 38

2.2.1. Introduction 38
2.2.2. Markov chain representation 40
2.2.3. One-sided control charts 41
2.2.3.1.Upper one-sided control charts 41
2.2.3.2.Lower one-sided control charts 49
2.2.4. Two-sided control charts 51
2.2.5. Summary 63

2.3. The tabular CUSUM control chart 64

2.3.1. Introduction 64
2.3.2. One-sided control charts 68
2.3.2.1.Upper one-sided control charts 68
2.3.2.2.Lower one-sided control charts 86
2.3.3. Two-sided control charts 91
2.3.4. Summary 106

2.4. The EWMA control chart 106

2.4.1. Introduction 106
2.4.2. The proposed EWMA sign chart 109
2.4.3. Markov-chain approach 111
2.4.4. Summary 120

3. Signed-rank control charts **121**

3.1. The Shewhart-type control chart 121

3.1.1. Introduction 121
3.1.2. Definition of the signed-rank test statistic 121
3.1.3. Plotting statistic 122
3.1.4. Determination of chart constants 125
3.1.5. Summary 129

3.2. The Shewhart-type control chart with runs-type signalling rules 129

3.2.1. Introduction 129
3.2.2. Example 130
3.2.3. Summary 132

3.3. The tabular CUSUM control chart 133

3.3.1. Introduction 133
3.3.2. One-sided control charts 133
3.3.2.1.Upper one-sided control charts 133
3.3.2.2.Lower one-sided control charts 161
3.3.3. Two-sided control charts 168
3.3.4. Summary 187

3.4. The EWMA control chart 188

3.4.1. Introduction 188
3.4.2. The proposed EWMA signed-rank chart 188
3.4.3. Markov-chain approach 189
3.4.4. Summary 196

Section B: Monitoring the location of a process when the target location is unspecified or unknown (Case U)

4. Sign-like control charts 197

4.1. The Shewhart-type control chart 197

4.1.1. Introduction 197
4.1.2. Preliminary 198
4.1.3. Probability of no signal 201
4.1.4. Determination of chart constants 202
4.1.5. The median chart 204
4.1.6. Control charts for other percentiles 205
4.1.7. Properties of order statistics 206
4.1.8. One-sided control charts 207
4.1.8.1. Lower one-sided control charts 207
4.1.8.2. Upper one-sided control charts 217
4.1.9. Two-sided control charts 227
4.1.10. Run-length distribution and *ARL* under some alternatives 238
4.1.10.1. Location alternatives 239
4.1.10.2. Scale alternatives 239
4.1.10.3. Location-scale alternatives 239
4.1.10.4. Lehmann alternatives 240
4.1.10.5. Proportional hazards alternatives 240
4.1.10.6. Summary 240

4.2. The Shewhart-type control chart with runs-type signalling rules 240

4.2.1. Introduction 240
4.2.2. Example 241
4.2.3. Summary 244

5. Signed-rank-like control charts **245**

5.1. The Shewhart-type control chart 245

5.1.1. Introduction 245
5.1.2. Definition of the signed-rank-like test statistic 245
5.1.3. Distribution-free properties 249
5.1.4. Simulation study 253
5.1.5. Comparisons 255
5.1.6. The tabular CUSUM control chart 263
5.1.7. The EWMA control chart 264
5.1.8. Summary 265

6. Mann-Whitney-Wilcoxon control charts **266**

6.1. The Shewhart-type control chart 266

6.1.1. Introduction 266
6.1.2. Plotting statistic 267
6.1.3. Properties of the run-length distribution 268
6.1.4. The computation of the signal probability 272
6.1.5. Saddlepoint approximations 273
6.1.6. Monte Carlo simulation 276
6.1.7. Determination of chart constants 282
6.1.8. Control chart performance 287
6.1.9. Summary 293

6.2. The tabular Phase I CUSUM control chart 294

6.2.1. Introduction 294
6.2.2. Determination of chart constants 297
6.2.3. Performance comparison 298
6.2.4. Summary 308

7. Concluding remarks **309**

Appendix A **311**

Appendix of theorems and proofs

Appendix B **325**

Appendix of computer programs

References **361**

Chapter 1: Introduction

1.1. Notation

SPC	Statistical process control
NSPC	Nonparametric statistical process control
pmf	Probability mass function
cdf	Cumulative distribution function
pgf	Probability generating function
mgf	Moment generating function
cgf	Cumulant generating function
n	Sample size
$X_1, X_2, ...X_n$	Random variables in a sample
$x_1, x_2, ...x_n$	Observations in a sample
θ_0	Target value / Known or specified in-control location parameter[1]
CUSUM	Cumulative sum
EWMA	Exponentially weighted moving average
ARL	Average run length
ARL_0	In-control average run length
ARL_δ	Out-of-control average run length
$SDRL$	Standard deviation of the run length
MRL	Median run length
UCL	Upper control limit
CL	Center line
LCL	Lower control limit
FAR	False alarm rate
FAP	False alarm probability
VSI	Variable sampling interval
FSI	Fixed sampling interval
a_U	Upper action limit / Upper control limit
w_U	Upper warning limit
w_L	Lower warning limit
a_L	Lower action limit / Lower control limit
TPM	Transition probability matrix
A	Absorbent
NA	Non-absorbent

[1] The location parameter could be the mean, median or some percentile of the distribution. When the underlying distribution is known to be highly skewed, the median or some percentile is preferred to the mean.

1.2. Distribution of chance causes

One of the main goals of statistical process control (SPC) is to distinguish between two sources of variability, namely common cause (chance cause) variability and assignable cause (special cause) variability. Common cause variability is an inherent or natural (random) variability that is present in any repetitive process, whereas assignable cause variability is a result of factors that are not solely random. In SPC, the pattern of chance causes is usually assumed to follow some parametric distribution (such as the normal). 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. However, the chance distribution is either unknown or far from being normal in many applications and consequently the performance of standard control charts is highly affected in such situations. Thus there is a need for some easy to use, flexible and robust control charts that do not require normality or any other specific parametric model assumption about the underlying chance distribution. Distribution-free or nonparametric control charts can serve this broader purpose. On this point see for example, Woodall and Montgomery (1999) and Woodall (2000). These researchers and others provide more than enough reasons for the development of nonparametric control charts.

1.3. Nonparametric or distribution-free

The term nonparametric is not intended to imply that there are no parameters involved, in fact, quite the contrary. While the term distribution-free seems to be a better description of what we expect from these charts, that is, they remain valid for a large class of distributions, nonparametric is perhaps the term more often used. In the statistics literature there is now a rather vast collection of nonparametric tests and confidence intervals and these methods have been shown to perform well compared to their normal theory counterparts. Remarkably, even when the underlying distribution is normal, the efficiency of some nonparametric methods relative to the corresponding (optimal) normal theory methods can be as high as 0.955 (see, e.g., Gibbons and Chakraborti, 2003). In fact, for some heavy-tailed distributions like the double exponential, nonparametric tests can be more efficient. It may be argued that nonparametric methods will be "less efficient" than their parametric counterparts when one has a complete knowledge of the process distribution for which that parametric method was specifically designed. However, the reality is that such information is seldom, if ever,

available in practice. Thus it seems natural to develop and use nonparametric methods in SPC and the quality practitioners will be well advised to have these techniques in their toolkits.

We only discuss univariate nonparametric control charts designed to track the location of a continuous process since very few charts are available for monitoring the scale and simultaneously monitoring the location and scale of a process. The field of multivariate control charts is interesting and the body of literature on nonparametric multivariate control charts is growing. However, in our opinion, it hasn't yet reached a critical mass and a discussion on this topic is better postponed for the future.

1.4. Nonparametric control charts

Chakraborti, Van der Laan and Bakir (2001) (hereafter CVB) provided a systematic and thorough account of the nonparametric control chart literature. A nonparametric control chart is defined in terms of its in-control run length distribution. If the in-control run length distribution of a control chart is the same for every continuous distribution, the chart is called nonparametric or distribution-free. CVB summarized the advantages of nonparametric control charts as follows: (i) simplicity, (ii) no need to assume a particular parametric distribution for the underlying process, (iii) the in-control run length distribution is the same for all continuous distributions, (iv) more robust and outlier resistant, (v) more efficiency in detecting changes when the true distribution is markedly non-normal, particularly with heavier tails, and (vi) no need to estimate the variance to set up charts for the location parameter. It is emphasized that from a technical point of view most nonparametric procedures require the population to be continuous in order to be distribution-free and thus in a SPC context we consider the so-called "variables control charts." Some disadvantages of nonparametric control charts are as follows: (i) they will be "less efficient" than their parametric counterparts when one has a complete knowledge of the process distribution for which that parametric method was specifically designed, (ii) one usually requires special tables when the sample sizes are small, and (iii) nonparametric methods are not well-known amongst all researchers and quality practitioners.

1.5. Terminology and formulation

Two important problems in usual SPC are monitoring the process mean and/or the process standard deviation. In the nonparametric setting, we consider, more generally, monitoring the center or the location (or a shift) parameter and/or a scale parameter of a process. The location parameter represents a typical value and could be the mean or the median or some percentile of the distribution; the latter two are especially attractive when the underlying distribution is expected to be skewed. Also in the nonparametric setting, the processes are implicitly assumed to follow (i) a location model, with a cdf $F(x-\theta)$, where θ is the location parameter or (ii) a scale model, with a cdf $F\left(\frac{x}{\tau}\right)$, where $\tau(>0)$ is the scale parameter. Even more generally, one might consider (iii) the location-scale model with cdf $F\left(\frac{x-\theta}{\tau}\right)$, where θ and τ are the location and the scale parameter, respectively. Under these model assumptions, the problem is to track θ and τ (or both), based on random samples or subgroups taken (usually) at equally spaced time points. In the usual (parametric) control charting problems F is assumed to be the cdf Φ of the standard normal distribution whereas in the nonparametric setting, for variables data, F is some unknown continuous cdf. Although the location-scale model seems to be a natural model to consider paralleling the normal theory case with mean and variance both unknown, most of what is currently available in the nonparametric statistical process control (NSPC) literature deals mainly with the location model.

As we noted earlier, a comprehensive survey of the literature until about 2000 can be found in CVB. Here, we mention some of the key contributions and ideas and a few of the more recent developments in the area; the literature on nonparametric methods continues to grow at a rapid pace. In fact, Woodall and Montgomery (1999) stated: 'There would appear to be an increasing role for nonparametric methods, particularly as data availability increases'. Most nonparametric charts, however, have been developed for Phase II applications. There are generally two phases in SPC. In Phase I (also called the retrospective phase), typically, preliminary analysis is done which includes planning, administration, data collection, data management, exploratory work including graphical and numerical analysis, goodness-of-fit analysis etc. to ensure that the process is in-control. This means that the process is managed to operate at or near some acceptable target value along with some natural variation and no

special causes of concern are expected to be present. Once this is ascertained, SPC moves to the next phase, Phase II, (or the prospective phase), where the control limits and/or the estimators obtained in Phase I are used for process monitoring based on new 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 parameters are unknown and need to be estimated, it is typically done in Phase I, with in-control data. This situation is referred to as the "standard(s) unknown" case and is denoted case U. In this text we are going to consider decision problems under both Phase I and Phase II. One of the main differences between the two phases is the fact that the *FAR* (or in-control average run length ARL_0) 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 many comparisons, whereas the *FAR* is the probability of a single false alarm involving only a single comparison. Various authors have studied the Phase I problem; see for example King (1954), Chou and Champ (1995), Sullivan and Woodall (1996), Jones and Champ (2002), Champ and Chou (2003), Champ and Jones (2004), Koning (2006) and Human, Chakraborti and Smit (2007). Since not much is typically known or can be assumed about the underlying process distribution in a Phase I setting, nonparametric Phase I control charts are of great use.

There are three main classes of control charts: the Shewhart chart, the cumulative sum (CUSUM) chart and the exponentially weighted moving average (EWMA) chart and their refinements. Relative advantages and disadvantages of these charts are well documented in the literature (see, e.g., Montgomery, 2001). Analogs of these charts have been considered in the nonparametric setting. We describe some of the charts under each of the three classes.

1.6. Shewhart-type 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. Both one-sided and two-sided charts are considered. The one-sided charts are more useful when only a directional shift (higher or lower) in the median is of interest. The two-sided charts, on the other hand, are typically used to detect a shift or change in the median in any direction.

1.7. CUSUM-type 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. These charts, typically based on the cumulative totals of a plotting statistic, obtained as data accumulate, are known to be more efficient for detecting certain types of shifts in the process. 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. It seems natural to consider analogs of these charts using the nonparametric plotting statistics discussed earlier. These lead to nonparametric CUSUM (NPCUSUM) charts.

1.8. EWMA-type charts

Another popular class of control charts is the exponentially weighted moving average (EWMA) 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 implement than the CUSUM charts.

Section A: Monitoring the location of a process when the target location is specified (Case K)

Chapter 2: Sign control charts

2.1. The Shewhart-type control chart

2.1.1. Introduction

The sign test is one of the simplest and broadly applicable nonparametric tests (see, e.g., Gibbons and Chakraborti, 2003) that can be used to test hypotheses (or find confidence intervals) for the median (or any specified percentile) of a continuous distribution. In this thesis, we will only consider the 50th percentile, i.e. the median. The fact that the sign test is applicable for any continuous population is an advantage to quality practitioners. Suppose that the median of a continuous process needs to be maintained at a specified value θ_0. Amin et al. (1995) presented Shewhart-type nonparametric charts for this problem using what are called "within group sign" statistics. This is called a sign chart (also referred to as the SN chart).

2.1.2. Definition of the sign test statistic

Let $X_{i1}, X_{i2}, ..., X_{in}$ denote the i^{th} $(i = 1,2,...)$ sample or subgroup of independent observations of size $n > 1$ from a process with an unknown continuous distribution function F. Let θ_0 denote the known or specified value of the median when the process is in-control, then θ_0 is called the target value. Compare each x_{ij} $(j = 1,2,...,n)$ with θ_0. Record the difference between θ_0 and each x_{ij} by subtracting θ_0 from x_{ij}. There will be n such differences, $x_{ij} - \theta_0$ $(j = 1,2,...,n)$, in the i^{th} sample. Let n^+ denote the number of observations with values greater than θ_0 in the i^{th} sample. Let n^- denote the number of observations with values less than θ_0 in the i^{th} sample. Provided there are no ties we have that $n^+ + n^- = n$.

Define

$$SN_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0) \tag{2.1}$$

where $sign(x) = -1, 0, 1$ if $x < 0, = 0, > 0$.

Then SN_i is the difference between n^+ and n^- in the i^{th} sample, i.e. SN_i is the difference between the number of observations with values greater than θ_0 and the number of observations with values less than θ_0 in the i^{th} sample.

Define

$$T_i = \frac{SN_i + n}{2}, \tag{2.2}$$

assuming there are no ties within a subgroup. The random variable T_i is the number of sample observations greater than or equal to θ_0 in the i^{th} sample. In (2.2) the statistic T_i is expressed in terms of the sign test statistic SN_i. Using the relationship in (2.2), the sign test statistic SN_i can be expressed in terms of the statistic T_i (if there are no ties within a subgroup) and we obtain

$$SN_i = 2T_i - n. \tag{2.3}$$

This relationship is evident from the fact that

$$SN_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0) = \sum_{j=1}^{n} \left(2\psi(x_{ij} - \theta_0) - 1\right) = 2T_i - n$$

where $\psi(x) = 0, 1$ if $x \leq 0, > 0$.

In the literature the statistic T_i is also well-known under the name *sign test statistic* (see, for example, Gibbons and Chakraborti (2003)). For the purpose of this study, SN_i will be referred to as the sign test statistic.

Zero differences

For a continuous random variable, X, the probability of any particular value is zero; thus, $P(X = a) = 0$ for any a. Since the distribution of the observations is assumed to be continuous, $P(X_{ij} - \theta_0 = 0) = 0$. Theoretically, the case where $sign(x_{ij} - \theta_0) = 0$ should occur with zero probability, but in practice zero differences do occur as a result of, for example, truncation or rounding of the observed values. A common practice in such cases is to discard all the observations leading to zero differences and to redefine *n* as the number of nonzero differences.

2.1.3. Plotting statistic

Sign control charts are based on the well-known sign test. A control chart is a graph consisting of values of a plotting (or charting) statistic and the associated control limits. The plotting statistic for the sign chart is $SN_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0)$ for $i = 1,2,3,\ldots$.

Distributional properties of the charting statistic

The random variable T_i has a binomial distribution with parameters n and $p = P(X_{ij} \geq \theta_0)$, i.e. $T_i \sim BIN(n, p)$. Hence, we can find the distribution of SN_i via the relationship given in (2.3).

Table 2.1. Moments and the probability mass function of the T_i and SN_i statistics, respectively.

	T_i	SN_i
Expected value	$E(T_i) = np$	$E(SN_i)$ $= E(2T_i - n)$ $= n(2p-1)$
Variance	$\text{var}(T_i) = np(1-p)$	$\text{var}(SN_i)$ $= \text{var}(2T_i - n)$ $= 4np(1-p)$
Standard deviation	$stdev(T_i) = \sqrt{np(1-p)}$	$stdev(SN_i)$ $= 2\sqrt{np(1-p)}$
Probability mass function (pmf)	$f(t) = P(T_i = t) = \binom{n}{t} p^t (1-p)^{n-t}$	$f(s)$ $= P(SN_i = s)$ $= P(2T_i - n = s)$ $= P\left(T_i = \frac{n+s}{2}\right)$

The probability distributions of T_i and SN_i are both **symmetric**[*] as long as the median remains at θ_0. In this case:

- the probability distributions are referred to as the *in-control* probability distributions;
- $p = P(X_{ij} \geq \theta_0) = 0.5$; and
- since the in-control distribution of the plotting statistic SN_i is symmetric, the control limits will be equal distances away from 0.

Figure 2.1 illustrates for $n = 10$ that the in-control probability distributions of T_i and SN_i are symmetric about their means, that is, T_i is symmetric around its mean given by $n \times p = 10 \times 0.5 = 5$ and SN_i is symmetric around its mean given by $n(2p-1) = 10(2 \times 0.5 - 1) = 0$.

[*] T_i and SN_i are symmetric about np and zero, respectively, as long as the median remains at θ_0.

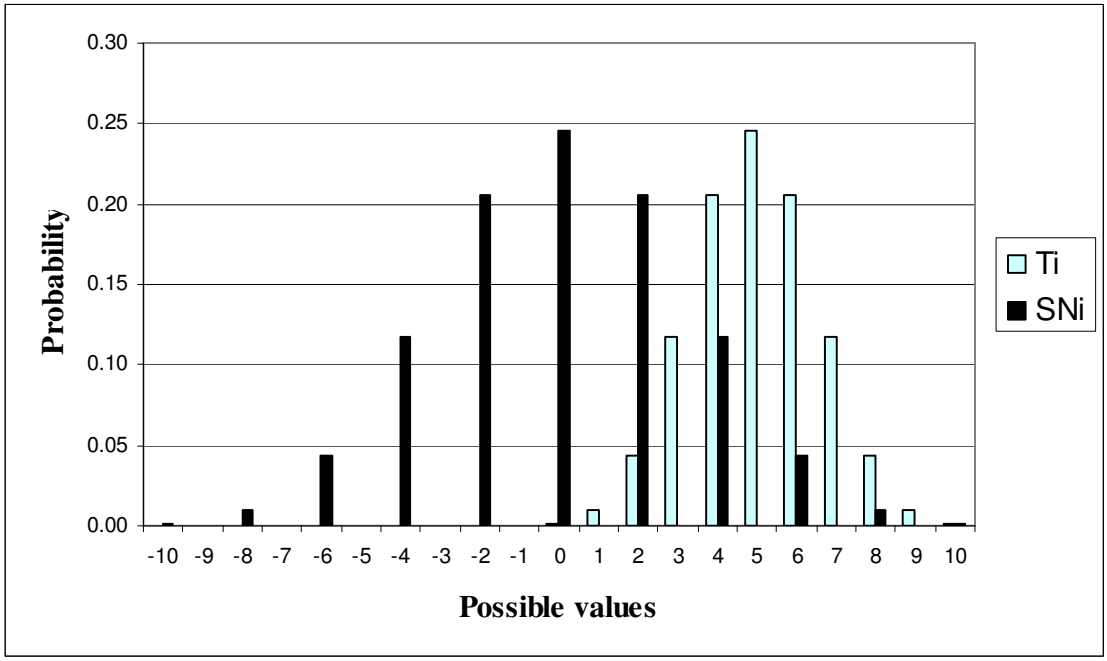

Figure 2.1. The in-control probability distribution of T_i and SN_i for $n = 10$.

2.1.4. Determination of control limits

In order to find the control limits and study chart performance, the distribution of SN_i is necessary; this can be most easily obtained using the relationship $SN_i = 2T_i - n$. Since the in-control distribution of T_i is binomial with parameters n and 0.5, it follows that the in-control distribution of SN_i is symmetric about 0 and hence the control limits and the center line of the two-sided nonparametric Shewhart-type sign chart (for the median) are given by

$$\begin{aligned} UCL &= c \\ CL &= 0 \\ LCL &= -c \end{aligned} \tag{2.4}$$

where $c \in \{1,2,...,n\}$.

If the plotting statistic SN_i falls between the control limits, that is, $-c < SN_i < c$, the process is declared to be in-control, whereas if the plotting statistic SN_i falls on or outside one of

the control limits, that is, if $SN_i \leq -c$ or $SN_i \geq c$, the process is declared to be out-of-control. In the latter case corrective action and a search for assignable causes is necessary.

Take note that T_i can also be calculated and plotted against the control limits. This is done by assuming that the LCL is equal to some constant a and that the UCL is equal to some constant b, i.e. the control limits are given by: $UCL = b$ and $LCL = a$. Since the in-control probability distribution of T_i is symmetric when working with the median, that is, $P(T_i = a) = P(T_i = n - a)$, a sensible choice for b is therefore $n - a$.

The control limits and the center line of the nonparametric Shewhart chart (for the median) using T_i as the plotting statistic are given by

$$UCL = n - a$$
$$CL = np$$
$$LCL = a$$

where a denotes a positive integer which is selected such that $LCL < UCL$.

Although both T_i and SN_i can be calculated for each sample and be compared to the control limits, the statistic SN_i has the advantage of keeping the control limits symmetric around zero. Therefore, the plotting statistic SN_i is calculated and used as the plotting statistic. The terms 'plotting statistic' and 'charting statistic' will be used interchangeably throughout this text.

The question arises: When using SN_i as the plotting statistic, what should the values of the control limits be set equal to? In other words, what is the value of the charting constant c? Specifying control limits is one of the critical decisions that must be made in designing a control chart. By moving the control limits farther away from the center line, we decrease the risk of a type I error – that is, the risk of a point falling beyond the control limits, indicating an out-of-control condition when no assignable cause is present. However, widening the control limits will also increase the risk of a type II error – that is, the risk of a point falling between the control limits when the process is really out-of-control. If we move the control limits closer to the center line, the opposite effect is obtained: The risk of type I error is increased, while the risk of type II

error is decreased. Consequently, the control limits are chosen such that if the process is in-control, nearly all of the sample points will fall between them. In other words, the charting constant c is typically obtained for a specified in-control average run length, which, in case K, is equal to the reciprocal of the nominal FAR, α. Thus, using the symmetry of the binomial distribution, c is the smallest integer such that $P_{\theta_0}(SN_i \geq c) \leq \frac{\alpha}{2}$. For example, using Table G of Gibbons and Chakraborti (2003) we give some $P_{\theta_0}(T \geq t)$ values in Table 2.2 that may be considered "small" in a SPC context. The charting constant c is obtained using $c = 2t - n$ (recall that the sign test statistic SN_i is expressed in terms of the statistic T_i by the relationship $SN_i = 2T_i - n$). The false alarm rate is obtained by adding the probability in the left tail, $P_{\theta_0}(T \leq n-t)$, and the probability in the right tail, $P_{\theta_0}(T \geq t)$, i.e. $FAR = P_{\theta_0}(T \leq n-t) + P_{\theta_0}(T \geq t)$. Since the probability distribution of T_i is symmetric (as long as the median remains at θ_0), the FAR is also obtained using $FAR = 2P_{\theta_0}(T \geq t)$. For example, for $n = 5$ we get $t = 5$ and thus $c = 5$ for a FAR of $2(0.0312) = 0.0624$ and this is the lowest FAR achievable. However for $n = 10$ the FAR drops to 0.0020 if $c = 10$. It should be noted that the lowest attainable FAR is always obtained when $n = t$.

Table 2.2. FAR and ARL_0 of a sign control chart for various values of $n = t$.

n	5	6	7	8	9	10
$P_{\theta_0}(T \geq t)$	0.0312	0.0156	0.0078	0.0039	0.0020	0.0010
$FAR\ (\alpha)$	0.0624	0.0312	0.0156	0.0078	0.0040	0.0020
ARL_0	16.00	32.00	64.00	128.00	256.00	512.00

Looking at the attainable FAR and ARL_0 values shown in Table 2.2, we see that unless the sample size is at least 10, the sign chart would be somewhat unattractive (from an operational point of view) in SPC applications, where one often stipulates a large in-control average run length, as large as 370 or 500, and a small FAR, as small as 0.0027. If, for example, the FAR is too 'large', which is the case for 'small' sample sizes, many false alarms will be expected by this chart leading to a possible loss of time and resources. Then again, the sign chart is the simplest of nonparametric charts that works under minimal assumptions. In fact, from the hypothesis testing

literature, it is known that the sign test (and so the chart) is more robust and efficient when the chance distribution is symmetric like the normal but with heavier tails such as the double exponential.

Example 2.1

A Shewhart-type sign chart for the Montgomery (2001) piston ring data

We illustrate the Shewhart-type sign chart using a set of data from Montgomery (2001; Tables 5.1 and 5.2) on the inside diameters of piston rings manufactured by a forging process. A part of this data, fifteen prospective samples (Table 5.2) each of five observations, is used here. The rest of the data (Table 5.1) will be used later. We assume that the underlying distribution is symmetric with a known median $\theta_0 = 74\,mm$. From Table G (see Gibbons and Chakraborti (2003)) we obtain $t = 5$ (when $n = 5$) for an achieved false alarm rate of $2(0.0312) = 0.0624$. Therefore, $c = 2\times 5 - 5 = 5$ and the control limits and the center line of the nonparametric Shewhart sign chart are given by $UCL = 5$, $CL = 0$ and $LCL = -5$.

Panel *a* of Table 2.3 displays the sample number. The two rows of each cell in panel *b* shows the individual observations and $sign(x_{ij} - \theta_0)$ values, respectively. The SN_i and T_i values are shown in panel *c* and panel *d*, respectively.

As an example, the calculation of SN_1 (found in Table 2.3) is given.

$$\begin{aligned} SN_1 &= sign(x_{11} - \theta_0) + sign(x_{12} - \theta_0) + sign(x_{13} - \theta_0) + sign(x_{14} - \theta_0) + sign(x_{15} - \theta_0) \\ &= sign(74.012 - 74) + sign(74.015 - 74) + sign(74.030 - 74) + sign(73.986 - 74) + sign(74 - 74) \\ &= sign(0.012) + sign(0.015) + sign(0.03) + sign(-0.014) + sign(0) \\ &= 1 + 1 + 1 - 1 + 0 \\ &= 2. \end{aligned}$$

Table 2.3. Data and calculations for the Shewhart sign chart[*].

Panel *a*	**Panel *b***					**Panel *c***	**Panel *d***
Sample number	**Individual observations** $sign(x_{ij}-\theta_0)$					SN_i	T_i
1	74.012[†] 1	74.015 1	74.030 1	73.986 -1	74.000 0	2	3
2	73.995 -1	74.010 1	73.990 -1	74.015 1	74.001 1	1	3
3	73.987 -1	73.999 -1	73.985 -1	74.000 0	73.990 -1	-4	0
4	74.008 1	74.010 1	74.003 1	73.991 -1	74.006 1	3	4
5	74.003 1	74.000 0	74.001 1	73.986 -1	73.997 -1	0	2
6	73.994 -1	74.003 1	74.015 1	74.020 1	74.004 1	3	4
7	74.008 1	74.002 1	74.018 1	73.995 -1	74.005 1	3	4
8	74.001 1	74.004 1	73.990 -1	73.996 -1	73.998 -1	-1	2
9	74.015 1	74.000 0	74.016 1	74.025 1	74.000 0	3	3
10	74.030 1	74.005 1	74.000 0	74.016 1	74.012 1	4	4
11	74.001 1	73.990 -1	73.995 -1	74.010 1	74.024 1	1	3
12	74.015 1	74.020 1	74.024 1	74.005 1	74.019 1	5	5
13	74.035 1	74.010 1	74.012 1	74.015 1	74.026 1	5	5
14	74.017 1	74.013 1	74.036 1	74.025 1	74.026 1	5	5
15	74.010 1	74.005 1	74.029 1	74.000 0	74.020 1	4	4

The sign chart is shown in Figure 2.2 with $UCL = 5$, $CL = 0$ and $LCL = -5$.

[*] See SAS Program 1 in Appendix B for the calculation of the values in Table 2.3.

[†] The two rows of each cell in panel *b* shows the x_{ij} and $sign(x_{ij}-\theta_0)$ values, respectively, for example,

$x_{11} = 74.012$ $sign(x_{11}-\theta_0) = 1$	is presented as	74.012 1

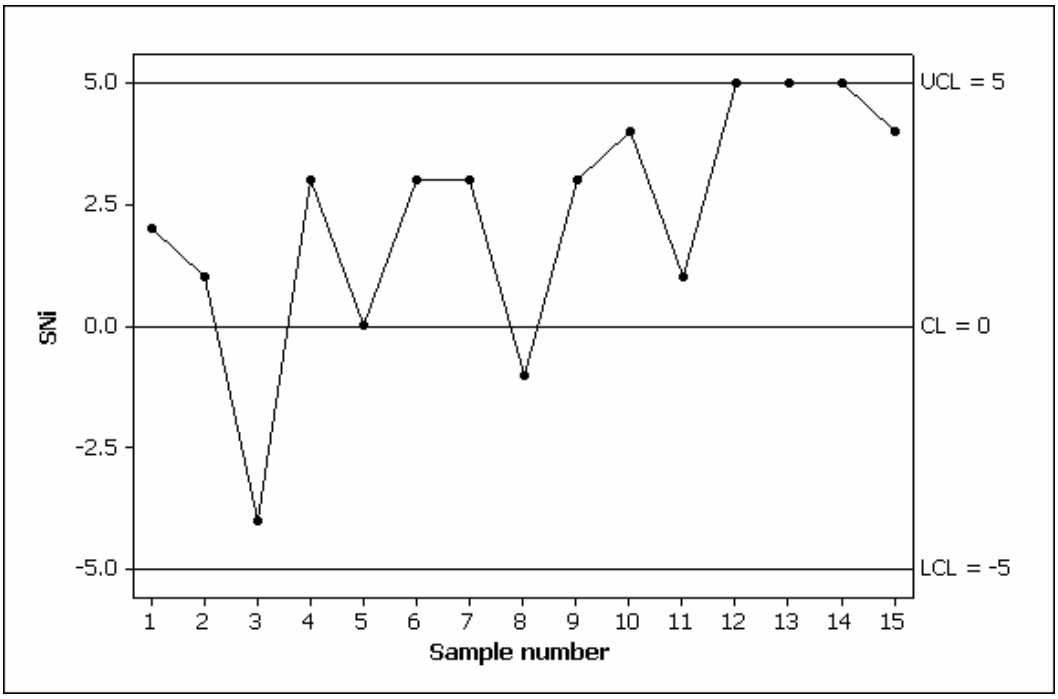

Figure 2.2. Shewhart-type sign control chart for Montgomery (2001) piston ring data.

Observations 12, 13 and 14 lie on the upper control limit which indicates that the process is out-of-control starting at sample 12. It appears most likely that the process median has shifted upwards from the target value of 74 *mm*. Corrective action and a search for assignable causes is necessary.

Control charts are often compared on the basis of various characteristics of the run length distribution, such as the ARL. One prefers a "large" in-control average run length (denoted ARL_0) and a "small" out-of-control ARL (denoted ARL_δ) under a shift. Amin et al. (1995) compared the ARL of the classical Shewhart $\overline{X}$ chart and the Shewhart-type sign chart for various shift sizes and underlying distributions. One practical advantage of sign charts, and of all nonparametric charts (if, of course, their assumptions are satisfied), is that the FAR (and the ARL_0) remains the same (eg. $FAR = 0.0624$ and $ARL_0 = 16$ for $n = 5$) for all continuous distributions. This is so because the in-control run length distribution is the same for every continuous distribution, for nonparametric charts, by definition. This does not hold for parametric charts (except for EWMA charts), and, as a result, parametric charts (again, with the EWMA chart being the exception) do not enjoy the same kind of robustness properties as nonparametric

charts do. It should be noted that the EWMA control chart can be designed so that it is robust to the normality assumption. On this point, Borror, Montgomery and Runger (1999) showed that the ARL_0 of the EWMA chart is reasonably close to the normal-theory value for both skewed and heavy-tailed symmetric non-normal distributions.

2.1.5. Run length distribution

The number of subgroups or samples that need to be collected (or, equivalently, the number of plotting statistics that must be plotted) before the next out-of-control signal is given by a chart is called the run length. The run length is a random variable denoted by N. A popular measure of chart performance is the 'expected value' or the 'mean' of the run length distribution, called the average run length (ARL). Various researchers, see for example, Barnard (1959) and Chakraborti (2007), have suggested using other characteristics for assessment of chart performance, for example, the standard deviation of the run length distribution ($SDRL$), the median run length (MRL) and/or other percentiles of the run length distribution. This recommendation is warranted seeing as (i) the run-length can only take on positive integer values by definition, (ii) the shape of this distribution is significantly right-skewed and (iii) it's known that in a right-skewed distribution the mean is greater than the median and thus is usually not a fair representation of a typical observation or the center.

Since the observations plotted on the control chart are assumed to be independent, the number of points that must be plotted until the first plotted point plots on or exceeds a control limit is a geometric random variable with parameter p, where p denotes the probability of a success (or, equivalently, the probability of a signal). Therefore, $N \sim GEO(P(\text{Signal}))$ where $P(\text{Signal}) = p$. The well-known properties of the geometric distribution are given in panel *a* of Table 2.4 and we use the fact that if q denotes the probability of no signal then $P(\text{Signal}) + P(\text{No Signal}) = p + q = 1$, i.e. $q = 1 - p$. The properties of the run length N are derived using the well-known properties of the geometric distribution and they are displayed in panel *b* of Table 2.4.

Table 2.4. The properties of the geometric and run length distribution.

	a	***b***
	X ~ GEO(p)	***N ~ GEO(P(Signal))***
Expected value	$E(X)=\frac{1}{p}$	$E(N)=ARL=\frac{1}{P(\text{Signal})}$
Variance	$\text{var}(X)=\frac{q}{p^2}$	$\text{var}(N)=\frac{1-P(\text{Signal})}{(P(\text{Signal}))^2}$
Standard deviation	$stdev(X)=\frac{\sqrt{q}}{p}$	$SDRL=\frac{\sqrt{1-P(\text{Signal})}}{P(\text{Signal})}$
Probability mass function (pmf)	$f(x)=P(X=x)=p(1-p)^{x-1}$ for $x=1,2,3,...$	$P(N=a)=P(\text{Signal})(1-P(\text{Signal}))^{a-1}$ for $a=1,2,3,...$
Cumulative distribution function (cdf)	$F(x)=P(X\le x)=1-(1-p)^x$ for $x=1,2,3,...$	$P(N\le a)=1-(1-P(\text{Signal}))^a$ for $a=1,2,3,...$

The $100\rho^{th}$ $(0<\rho<1)$ percentile is defined as the smallest l such that the cdf, given by $P(N\le l)=1-(1-P(\text{Signal})^l$ for $l=1,2,.....$, at the integer l is at least $(100\times\rho)\%$, that is,

$$l=\min\{j:1-(1-P(\text{Signal}))^j\ge\rho\} \text{ for } j=1,2,... \tag{2.5}$$

which reduces to finding the smallest positive integer l such that

$$l\ge\frac{\ln(1-\rho)}{\ln(1-P(\text{Signal}))}. \tag{2.6}$$

The run length distribution can be described via percentiles, for example, using the 5th, 25th (the first quartile, Q_1), 50th (the median run length, MRL), 75th (the third quartile, Q_3) and the 95th percentiles by substituting ρ in expression (2.6) by 0.05, 0.25, 0.50, 0.75 and 0.95, respectively.

2.1.6. One-sided control charts

A lower one-sided chart will have a LCL equal to some constant value with no UCL, whereas an upper one-sided chart will have an UCL equal to some constant value with no LCL. One-sided control charts are particularly useful in situations where only an upward (or only a downward) shift in a particular process parameter is of interest. For example, we might be monitoring the breaking strength of material used to make parachutes. If the breaking strength of the material decreases it might tear at a critical time, whereas if the breaking strength of the material increases it is beneficial to the user, since the material would, most likely, not tear while being used. In such a scenario a lower one-sided chart will be sufficient, since we are only interested in detecting a downward shift in a process parameter.

For the sign control chart, if we are only interested in detecting a downward shift we will use a lower one-sided sign control chart with $LCL = -c$ and no upper control limit. Consequently, if the plotting statistic SN_i falls on or below the LCL the process is declared to be out-of-control. On the other hand, if we are only interested in detecting an upward shift we will use an upper one-sided sign control chart with $UCL = c$ and no lower control limit. Consequently, if the plotting statistic SN_i falls on or above the UCL the process is declared to be out-of-control.

2.1.6.1. Lower one-sided control charts

Result 2.1: Probability of a signal

The probability that the control chart signals, that is, the probability that the plotting statistic SN_i is smaller than or equal to the lower control limit, can be expressed in terms of $p = P(X_{ij} \geq \theta_0)$, the sample size n and the constant c. Let $P^L(\text{Signal})$ denote the probability of a signal, where superscript L refers to the lower one-sided chart. The probability of a signal is then given by

$$P^L(\text{Signal}) = P^L(SN_i \le LCL) = P^L(SN_i \le -c) = P^L(2T_i - n \le -c) = P^L\left(T_i \le \frac{n-c}{2}\right). \qquad (2.7)$$

Note that (2.7) can be solved by using the cdf of a Binomial distribution.

The probabilities, $P^L(\text{Signal})$'s, were computed using Mathcad (see Mathcad Program 2 in Appendix B). In doing so, we kept in mind that we ultimately wanted to use the T_i statistic in the calculation of $P^L(\text{Signal})$, because then we could use the cdf of a Binomial distribution to find $P^L(\text{Signal})$. Therefore, the probability of a signal for the lower one-sided sign chart was computed using

$$P^L(\text{Signal}) = P^L(T_i \le LCL) = P^L(T_i \le a) = \sum_{i=0}^{a}\binom{n}{i}p^i(1-p)^{n-i}. \qquad (2.8)$$

The results are given in Tables 2.5, 2.6 and 2.7 for $n = 5$, $n = 10$ and $n = 15$, respectively, for $p = 0.1(0.1)0.9$ and $a = 0(1)n$. The shaded column ($p = 0.5$) contains the value of the *in-control* average run length (ARL_0) and the false alarm rate (FAR), whereas the rest of the columns ($p \neq 0.5$) contain the values of the *out-of-control* average run length (ARL_δ) and the probability of a signal (when the process is considered to be out-of-control).

Result 2.2: Average run length

Since the run length has a geometric distribution (recall that $N \sim GEO(P^L(\text{Signal}))$ the expected value of this specific geometric distribution will be equal to $\frac{1}{P^L(\text{Signal})}$. The ARL is the mean of the run length distribution. Therefore, we have that

$$ARL^L = E(N) = \frac{1}{P^L(\text{Signal})}. \qquad (2.9)$$

Result 2.3: Standard deviation of the run length

Since the run length has a geometric distribution (see Result 2.2) the standard deviation will be equal to $\frac{\sqrt{1-P^L(\text{Signal})}}{P^L(\text{Signal})}$. Therefore, we have that

$$SDRL^L(N) = \frac{\sqrt{1-P^L(\text{Signal})}}{P^L(\text{Signal})}. \qquad (2.10)$$

Example 2.2

For a sample size of 10 ($n = 10$), $p = 0.5$ and $a = 2$, we can calculate the probability of a signal and the average run length using (2.8) and (2.9), respectively, and we obtain $P^L(\text{Signal}) = \sum_{i=0}^{2}\binom{10}{i}(0.5)^i(1-0.5)^{10-i} = 0.055$ and $ARL^L = \frac{1}{0.055} = 18.29$.

2.1.6.2. Upper one-sided control charts

Result 2.4: Probability of a signal

The probability that the control chart signals, that is, the probability that the plotting statistic SN_i is greater than or equal to the upper control limit, can be expressed in terms of $p = P(X_{ij} \geq \theta_0)$, the sample size n and the constant c. Let $P^U(\text{Signal})$ denote the probability of a signal, where superscript U refers to the upper one-sided chart. The probability of a signal is then given by

$$P^U(\text{Signal}) = P^U(SN_i \geq UCL) = P^U(SN_i \geq c) = P^U\left(T_i \geq \frac{n+c}{2}\right) = 1 - P^U\left(T_i \leq \frac{n+c}{2} - 1\right). \qquad (2.11)$$

Note that (2.11) can be solved by using the cdf of a Binomial distribution.

The probabilities, $P^U(\text{Signal})$'s, were computed using Mathcad (see Mathcad Program 1 in Appendix B). In doing so, we kept in mind that we ultimately wanted to use the T_i statistic in

the calculation of $P^U(\text{Signal})$, because then we could use the cdf of a Binomial distribution to find $P^U(\text{Signal})$. Therefore, the probability of a signal for the upper one-sided sign chart was computed using

$$P^U(\text{Signal}) = P^U(T_i \geq UCL) = P^U(T_i \geq n-a) = \sum_{i=n-a}^{n} \binom{n}{i} p^i (1-p)^{n-i}. \qquad (2.12)$$

The results are given in Tables 2.5, 2.6 and 2.7 for $n=5$, $n=10$ and $n=15$, respectively, for $p = 0.1(0.1)0.9$ and $a = 0(1)n$. The shaded column ($p=0.5$) contains the value of the *in-control* average run length (ARL_0) and the false alarm rate (FAR), whereas the rest of the columns ($p \neq 0.5$) contain the values of the *out-of-control* average run length (ARL_δ) and the probability of a signal (when the process is considered to be out-of-control).

Result 2.5: Average run length

Since the run length has a geometric distribution (recall that $N \sim GEO(P^U(\text{Signal}))$ the expected value of this specific geometric distribution will be equal to $\frac{1}{P^U(\text{Signal})}$. The ARL is the mean of the run length distribution. Therefore, we have that

$$ARL^U = E(N) = \frac{1}{P^U(\text{Signal})}. \qquad (2.13)$$

Result 2.6: Standard deviation of the run length

Since the run length has a geometric distribution (see Result 2.5) the standard deviation will be equal to $\frac{\sqrt{1-P^U(\text{Signal})}}{P^U(\text{Signal})}$. Therefore, we have that

$$SDRL^U(N) = \frac{\sqrt{1-P^U(\text{Signal})}}{P^U(\text{Signal})}. \qquad (2.14)$$

Example 2.3

For a sample size of 10 ($n = 10$), $p = 0.5$ and $a = 2$, we can calculate the probability of a signal and the average run length using (2.12) and (2.13), respectively, and we obtain $P^U(\text{Signal}) = \sum_{i=10-2}^{10} \binom{10}{i} (0.5)^i (1-0.5)^{10-i} = 0.055$ and $ARL^U = \frac{1}{0.055} = 18.29$.

Application

The average run length values for the lower and upper one-sided Shewhart sign charts are calculated by evaluating expressions (2.8) and (2.9) for the lower one-sided chart and expressions (2.12) and (2.13) for the upper one-sided chart using $n = 5, 10$ and 15, respectively. These values are shown in Table 2.5, Table 2.6 and Table 2.7, respectively.* As mentioned previously, the shaded column ($p = 0.5$) contains the value of the *in-control* average run length (ARL_0) and the false alarm rate (FAR), whereas the rest of the columns ($p \neq 0.5$) contain the values of the *out-of-control* average run length (ARL_δ) and the probability of a signal (when the process is considered to be out-of-control).

* Table 2.5, Table 2.6 and Table 2.7 should preferably be studied together.

Table 2.5. The average run length for the one-sided Shewhart sign chart with $n = 5$.[**]

ARL **P(Signal)**	p^U	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
	p^L	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
a	0	1.69 0.590	3.05 0.328	5.95 0.168	12.86 0.078	32.00 0.031	97.66 0.010	411.52 0.002	3125.00 0.000	100000.00 0.000
	1	1.09 0.919	1.36 0.737	1.89 0.528	2.97 0.337	5.33 0.188	11.49 0.087	32.49 0.031	148.81 0.007	2173.91 0.000
	2	1.01 0.991	1.06 0.942	1.19 0.837	1.47 0.683	2.00 0.500	3.15 0.317	6.13 0.163	17.27 0.058	116.82 0.009
	3	1.00 1.000	1.01 0.993	1.03 0.969	1.10 0.913	1.23 0.813	1.51 0.663	2.12 0.472	3.81 0.263	12.28 0.081
	4	1.00 1.000	1.00 1.000	1.00 0.998	1.01 0.990	1.03 0.969	1.08 0.922	1.20 0.832	1.49 0.672	2.44 0.410
	5	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000

[**] See Mathcad Programs 1 and 2 in Appendix B for the calculation of the values in Table 2.5.

Table 2.6. The average run length for the one-sided Shewhart sign chart with $n = 10$.[††]

ARL **P(Signal)**	p^U	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
	p^L	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
a	0	2.87 0.349	9.31 0.107	35.40 0.028	165.38 0.006	1024.00 0.001	9536.74 0.000	169350.88 0.000	9765625.00 0.000	10000000000.00 0.000
	1	1.36 0.736	2.66 0.376	6.70 0.149	21.57 0.046	93.09 0.011	596.05 0.002	6959.63 0.000	238185.98 0.000	109890109.89 0.000
	2	1.08 0.930	1.48 0.678	2.61 0.383	5.98 0.167	18.29 0.055	81.34 0.012	628.78 0.002	12832.62 0.000	2676659.53 0.000
	3	1.00 0.987	1.14 0.879	1.54 0.650	2.62 0.382	5.82 0.172	18.26 0.055	94.41 0.011	1156.93 0.001	109629.89 0.000
	4	1.00 0.998	1.03 0.967	1.18 0.850	1.58 0.633	2.65 0.377	6.02 0.166	21.12 0.047	157.00 0.006	6807.23 0.000
	5	1.00 1.000	1.01 0.994	1.05 0.953	1.20 0.834	1.61 0.623	2.73 0.367	6.65 0.150	30.49 0.033	611.64 0.002
	6	1.00 1.000	1.00 0.999	1.01 0.989	1.06 0.945	1.21 0.828	1.62 0.618	2.85 0.350	8.27 0.121	78.15 0.013
	7	1.00 1.000	1.00 1.000	1.00 0.998	1.01 0.988	1.06 0.945	1.20 0.833	1.62 0.617	3.10 0.322	14.25 0.070
	8	1.00 1.000	1.00 1.000	1.00 1.000	1.00 0.998	1.01 0.989	1.05 0.954	1.18 0.851	1.60 0.624	3.79 0.264
	9	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 0.999	1.01 0.994	1.03 0.972	1.12 0.893	1.54 0.651
	10	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000

[††] See Mathcad Programs 1 and 2 in Appendix B for the calculation of the values in Table 2.6.

Table 2.7. The average run length for the one-sided Shewhart sign chart with $n = 15$.[‡‡]

ARL P(Signal)	p^U	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
	p^L	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
a	0	4.86 0.206	28.42 0.035	210.63 0.005	2126.82 0.000	32768.00 0.000	931322.57 0.000	69691719.38 0.000	30517578125.00 0.000	999999999999987.00 0.000
	1	1.82 0.549	5.98 0.167	28.35 0.035	193.35 0.005	2048.00 0.000	39630.75 0.000	1935881.09 0.000	500288165.98 0.000	7352941176470.50 0.000
	2	1.23 0.816	2.51 0.398	7.88 0.127	36.88 0.027	270.81 0.004	3585.46 0.000	114687.42 0.000	17528764.00 0.000	115727346371.95 0.000
	3	1.06 0.944	1.54 0.648	3.37 0.297	11.05 0.091	56.89 0.018	518.73 0.002	10910.04 0.000	988871.98 0.000	2938272765.74 0.000
	4	1.01 0.987	1.20 0.836	1.94 0.515	4.60 0.217	16.88 0.059	106.98 0.009	1487.58 0.001	80245.85 0.000	107571980.98 0.000
	5	1.00 0.998	1.07 0.939	1.39 0.722	2.48 0.403	6.63 0.151	29.56 0.034	273.78 0.004	8831.92 0.000	5358475.36 0.000
	6	1.00 1.000	1.02 0.982	1.15 0.869	1.64 0.610	3.29 0.304	10.52 0.095	65.61 0.015	1273.91 0.001	351310.79 0.000
	7	1.00 1.000	1.00 0.996	1.05 0.950	1.27 0.787	2.00 0.500	4.69 0.213	19.99 0.050	235.86 0.004	29739.88 0.000
	8	1.00 1.000	1.00 0.999	1.02 0.985	1.11 0.905	1.44 0.696	2.56 0.390	7.63 0.131	55.37 0.018	3219.26 0.000
	9	1.00 1.000	1.00 1.000	1.00 0.996	1.04 0.966	1.18 0.849	1.68 0.597	3.59 0.278	16.38 0.061	444.51 0.002
	10	1.00 1.000	1.00 1.000	1.00 0.999	1.01 0.991	1.06 0.941	1.28 0.783	2.06 0.485	6.09 0.164	78.61 0.013
	11	1.00 1.000	1.00 1.000	1.00 1.000	1.00 0.998	1.02 0.982	1.10 0.909	1.42 0.703	2.84 0.352	18.00 0.056
	12	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 0.996	1.03 0.973	1.15 0.873	1.66 0.602	5.43 0.184
	13	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.01 0.995	1.04 0.965	1.20 0.833	2.22 0.451
	14	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 0.995	1.04 0.965	1.26 0.794
	15	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000	1.00 1.000

[‡‡] See Mathcad Programs 1 and 2 in Appendix B for the calculation of the values in Table 2.7.

2.1.7. Two-sided control charts

Result 2.7: Probability of a signal

The probability that the control chart signals, that is, the probability that the plotting statistic SN_i is greater than or equal to the UCL, or smaller than or equal to the LCL, can be expressed in terms of $p = P(X_{ij} \geq \theta_0)$, the sample size n and the constant c. The probability of a signal is then given by

$$P(\text{Signal}) = 1 - P(-c < 2T_i - n < c) = 1 - P\left(T_i \leq \frac{n+c}{2} - 1\right) + P\left(T_i \leq \frac{n-c}{2}\right). \quad (2.15)$$

Note that (2.15) can be solved by using the cdf of a binomial distribution.

Result 2.8: Average run length

Since the run length has a geometric distribution we have that

$$ARL = E(N) = \frac{1}{P(\text{Signal})}. \quad (2.16)$$

Compare expression (2.16) to expressions (2.9) and (2.13).

Result 2.9: Standard deviation of the run length

Since the run length has a geometric distribution we have that

$$SDRL(N) = \frac{\sqrt{1 - P(\text{Signal})}}{P(\text{Signal})}. \quad (2.17)$$

Compare expression (2.17) to expressions (2.10) and (2.14).

2.1.8. Summary

In Section 2.1 we have described and evaluated the nonparametric Shewhart-type sign control chart. Generally speaking, when the underlying process distribution is either asymmetric or symmetric with heavy tails, sign charts are more efficient while the reverse is true for normal and normal-like distributions with light tails. One practical advantage of the nonparametric Shewhart-type sign control chart is that there is no need to assume a particular parametric distribution for the underlying process (see Section 1.4 for other advantages of nonparametric charts).

2.2. The Shewhart-type control chart with warning limits

2.2.1. Introduction

It is known that standard Shewhart charts are efficient in detecting large process shifts quickly, but are insensitive to small shifts (see, for example, Montgomery (2005)). Additional supplementary rules have been suggested to increase the sensitivity of standard Shewhart charts to small process shifts. Shewhart (1941) gave the first proposal in making the standard Shewhart chart more sensitive to small process shifts by proposing that additional sensitizing tests should be incorporated into the standard Shewhart chart. Various rules or 'tests for special causes' have been considered in the literature for parametric control charts; see for example, the rules associated with the Shewhart control chart in Nelson (1984) and in the Western Electric handbook (1956). See also the discussion in Montgomery (2001).

Runs rules can be used to increase the sensitivity of standard Shewhart charts. Denote each runs rule by $R(r,n,k,l)$ where a signal is given if r out of the last n points fall in the interval (k,l), where $r \le n$ are integers and $k < l$. The well-known standardized Shewhart $\overline{X}$ control chart is denoted by $\{R(1,1,-\infty,-3) \cup R(1,1,3,\infty)\}$, since the standardized Shewhart $\overline{X}$ control chart signals if any charting statistic (1 out of 1 point) falls in the interval $(-\infty,-3)$ or if any charting statistic (1 out of 1 point) falls in the interval $(3,\infty)$.

Page (1955) considered a Markov-chain approach for simple combinations of runs rules. Amin et al. (1995) considered Shewhart-type sign charts with warning limits and runs rules. Page (1962), Weindling, Littauer and Oliveira (1979) and Champ and Woodall (1987) studied the properties of $\overline{X}$ charts with warning limits.

Incorporating the runs rules $\{R(1,1,-\infty,a_L)\cup R(1,1,a_U,\infty)\}$ into the Shewhart sign chart is similar to using *action limits* where action will be taken if any 1 point falls outside the action limits. Incorporating the two runs rules $\{R(r,r,w_U,a_U)\cup R(r,r,a_L,w_L)\}$ into the Shewhart sign chart is similar to using *warning limits* where action will be taken if r successive points fall between the warning and action limits, that is, action will be taken if r successive points fall between w_U and a_U *or* action will be taken if r successive points fall between a_L and w_L. Hence, rule A follows: Action will be taken if r successive points fall between w_U and a_U, *or* if r successive points fall between a_L and w_L, *or* if any point falls outside the action limits. Let L denote the ARL of rule A. Assume that the upper action and upper warning limits are equal to some constants represented by a and w, respectively, that is, $a_U = a$ and $w_U = w$. In the case of the Shewhart-type sign control chart with warning limits, sensible choices for the lower action and lower warning limits are $-a$ and $-w$, respectively, that is, $a_L = -a$ and $w_L = -w$. The latter choices are sensible, since the in-control distribution of SN_i is symmetric about zero (see Section 2.1.3).

In Section 2.2.3.1 two runs rules are incorporated into the upper one-sided Shewhart sign chart. Similarly, in Section 2.2.3.2 two runs rules are incorporated into the lower one-sided Shewhart sign chart. The average run lengths are computed for the upper and lower charts, respectively. Finally, in Section 2.2.4 two runs rules are incorporated into the two-sided Shewhart sign chart.

2.2.2. Markov chain representation

A Markov chain representation of a Shewhart chart supplemented with runs rules is used for calculating the probability that any subset of runs rules will give an out-of-control signal. In this section some basic concepts of matrices and transition probabilities are given and explained. An illustrative example follows in the next section, i.e. Section 2.2.3.1.

Let p_{ij} represent the probability that the process will, when in state i, next make a transition to state j. Since probabilities are non-negative, $p_{ij} \geq 0$. Let TPM denote the matrix of one-step transition probabilities. The abbreviation TPM will be used throughout the text for *transition probability matrix* which is given by

$$TPM_{(n+1)\times(n+1)} = [p_{ij}] = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0n} \\ p_{10} & p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{i0} & p_{i1} & \cdots & p_{in} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n0} & p_{n1} & \cdots & p_{nn} \end{pmatrix} \text{ for } i = 0,1,...,n \text{ and } j = 0,1,...,n \text{ with } n \geq 2.$$

The i^{th} row, $(p_{i0}, p_{i1},..., p_{in})$, contains all the transition probabilities to go from state i to one of the states in Ω, where Ω denotes the state space, i.e. $\Omega = \{0,1,...,n\}$. We have that

$$\sum_{j\in\Omega} p_{ij} = 1 \quad \forall i \tag{2.18}$$

since it's certain that starting in state i the process will go to one of the states in one step.

2.2.3. One-sided control charts

2.2.3.1. Upper one-sided control charts

The upper one-sided Shewhart sign chart, described previously, is efficient in detecting large process shifts quickly. Since it is known to be inefficient in detecting small process shifts, an upper warning limit is drawn below the upper action limit to increase its sensitivity for detecting small shifts.

Define rule A_U as: 'Action will be taken if r successive points fall between w_U and a_U (denoted by $R(r, r, w_U, a_U)$) *or* if any point falls above a_U (denoted by $R(1,1, a_U, \infty)$)'. Clearly, rule A_U is created to detect upward shifts. Let L^U denote the ARL of rule A_U. L^U can be calculated by enumerating the possible combinations of the positions of the plotted points and treating them as the states of a discrete Markov process. The following set of rules is used: $\{ R(r, r, w_u, a_u) \cup R(1,1, a_u, \infty) \}$. The 3 mutually exclusive intervals (also referred to as zones) which are considered are given by:

Zone Z_0 = the interval $(-\infty, w_U)$

Zone Z_1 = the interval $[w_U, a_U)$

Zone Z_2 = the interval $[a_U, \infty)$

These zones are graphically represented in Figure 2.3.

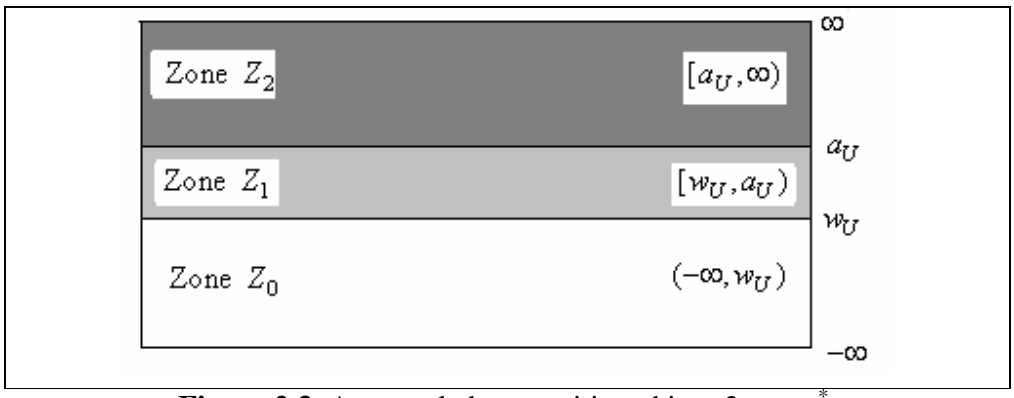

Figure 2.3. A control chart partitioned into 3 zones*.

$R(r, r, w_u, a_u)$: The chart will signal if any r successive points fall in Zone Z_1; or

$R(1,1, a_u, \infty)$: The chart will signal if any 1 point falls in Zone Z_2.

Classification of states

If a state is entered once and can't be left, the state is said to be absorbent. As a result, the probability of going from an absorbent state to the same absorbent state is equal to one. The transient (non-absorbent) states are the remaining states of which the time of return or the number of steps before return is uncertain.

Table 2.8. Classifications and descriptions of states.

State number	Description of state	Absorbent (A)/ Non-absorbent (NA)
0	1 point plots in Zone Z_0	NA
1	1 point plots in Zone Z_1	NA
2	2 successive points plot in Zone Z_1	NA
3	3 successive points plot in Zone Z_1	NA
⋮	⋮	⋮
$r-1$	$r-1$ successive points plot in Zone Z_1	NA
r	r successive points plot in Zone Z_1 or 1 point plots in Zone Z_2	A

* Any point plotting on a line is to be taken as plotting into the adjacent more extreme zone of the chart.

Let p_i denote the probability of plotting in Zone Z_i for $i = 0,1,2$. Therefore:

> p_0 is the probability of plotting in Zone Z_0; $p_0 = P(SN_i < w_U)$;
>
> p_1 is the probability of plotting in Zone Z_1; $p_1 = P(w_U \le SN_i < a_U)$; and
>
> p_2 is the probability of plotting in Zone Z_2; $p_2 = P(SN_i \ge a_U)$.

Clearly, $\sum_{i=0}^{2} p_i = p_0 + p_1 + p_2 = 1$, since the statistic **must** plot in one of the 3 zones. The transition probability matrix, $TPM = [p_{ij}]$, for $i = 0,1,2,...,r$ and $j = 0,1,2,...,r$ is given by

$$TPM_{(r+1)\times(r+1)} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0(r-1)} & p_{0r} \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1(r-1)} & p_{1r} \\ p_{20} & p_{21} & p_{22} & \cdots & p_{2(r-1)} & p_{2r} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ p_{(r-1)0} & p_{(r-1)1} & p_{(r-1)2} & \cdots & p_{(r-1)(r-1)} & p_{(r-1)r} \\ p_{r0} & p_{r1} & p_{r2} & \cdots & p_{r(r-1)} & p_{rr} \end{pmatrix} = \begin{pmatrix} p_0 & p_1 & 0 & \cdots & 0 & p_2 \\ p_0 & 0 & p_1 & \cdots & 0 & p_2 \\ p_0 & 0 & 0 & \cdots & 0 & p_2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ p_0 & 0 & 0 & \cdots & 0 & p_1 + p_2 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

From expression (2.18) we have $\sum_{j\in\Omega} p_{ij} = 1 \quad \forall i$. This is easily proven for $i = 0,1,2,...,r$.

For example, for $i = 0$ we have that $\sum_{j=0}^{r} p_{0j} = p_0 + p_1 + 0 + \cdots + 0 + p_2 = 1$. The rest of the calculations follow similarly. Table 2.9 illustrates that the TPM can be partitioned into 4 sections.

Table 2.9. Transition probabilities for a Markov chain with one absorbing state.

	States at time $t+1$					
States at time t	0 (NA)	1 (NA)	2 (NA)	…	$r-1$ (NA)	r (A)
0 (NA)	p_0	p_1	0	…	0	p_2
1 (NA)	p_0	0	p_1	…	0	p_2
2 (NA)	p_0	0	0	…	0	p_2
⋮	⋮	⋮	⋮	…	⋮	⋮
$r-1$ (NA)	p_0	0	0	…	0	$p_1 + p_2$
r (A)	0	0	0	…	0	1

=

States at time $t+1$					
0 (NA)	1 (NA)	2 (NA)	…	$r-1$ (NA)	r (A)
$Q_{r\times r}$					$\underline{p}_{r\times 1}$
$\underline{0}'_{1\times r}$					$1_{1\times 1}$

where the sub-matrix $Q_{r\times r} = \begin{pmatrix} p_0 & p_1 & 0 & \cdots & 0 \\ p_0 & & p_1 & \cdots & 0 \\ p_0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ p_0 & 0 & 0 & \cdots & 0 \end{pmatrix}$ is called the essential transition probability sub-matrix and it contains all the transition probabilities of going from a non-absorbent (transient) state to a non-absorbent state, $Q:(NA \to NA)$. $\underline{p}_{r\times 1} = (p_2 \quad p_2 \quad p_2 \quad \cdots \quad p_1 + p_2)'$ contains all the transition probabilities of going from each non-absorbent state to the absorbent states, $\underline{p}:(NA \to A)$. $\underline{0}'_{1\times r} = (0 \quad 0 \quad 0 \quad \cdots \quad 0)$ contains all the transition probabilities of going from each absorbent state to the non-absorbent states, $\underline{0}':(A \to NA)$. $\underline{0}'$ is a row vector with all its elements equal to zero, since it is impossible to go from an absorbent state to a non-absorbent state, because once an absorbent state is entered, it is never left. $1_{1\times 1}$ represents the scalar value one which is the probability of going from an absorbent state to an absorbent state, $1:(A \to A)$. Therefore,

$$TPM_{(r+1)\times(r+1)} = \begin{pmatrix} Q_{r\times r} & | & \underline{p}_{r\times 1} \\ - & - & - \\ \underline{0}'_{1\times r} & | & 1_{1\times 1} \end{pmatrix}. \tag{2.19}$$

Let L_i^U denote the run length of the upper one-sided chart with initial state i for $i = 0,1,2,...,r-1$. To calculate the probability mass function, define the $r\times 1$ vector $\underline{L}_h^U$ by

$$\underline{L}_h^U = \begin{pmatrix} P(L_0^U = h) \\ P(L_1^U = h) \\ \vdots \\ P(L_{r-1}^U = h) \end{pmatrix} = \begin{pmatrix} \text{Probability that the run length of the chart with initial state 0 is } h \\ \text{Probability that the run length of the chart with initial state 1 is } h \\ \vdots \\ \text{Probability that the run length of the chart with initial state } (r-1) \text{ is } h \end{pmatrix}.$$

Brook and Evans (1972) showed that these vectors can be calculated recursively using

$$\underline{L}_1^U = (I - Q)\underline{1}$$

and (2.20)

$$\underline{L}_h^U = Q\underline{L}_{h-1}^U \text{ for } h = 2,3,...$$

where $\underline{1}$ is a $r\times 1$ column vector of $1's$, I is the $r\times r$ identity matrix and Q is the $r\times r$ essential transition probability sub-matrix obtained from the partitioned TPM.

Multiplying out the matrices in (2.20) we get expressions for L_i^U for $i=1,2,...r-1$.

$$L_i^U = 1 + p_0 L_0^U + p_1 L_{i+1}^U \quad \text{for} \quad (i = 0,1,2,...,r-2)$$

and (2.21)

$$L_{r-1}^U = 1 + p_0 L_0^U .$$

These equations may be solved recursively for $L_1^U,...,L_{r-1}^U$ in terms of L_0^U:

$$L_0^U = 1 + p_0 L_0^U + p_1 L_1^U$$
$$L_0^U = 1 + p_0 L_0^U + p_1 (1 + p_0 L_0^U + p_1 L_2^U)$$
$$L_0^U = 1 + p_0 L_0^U + p_1 + p_1 p_0 L_0^U + p_1^2 L_2^U$$
$$L_0^U = 1 + p_0 L_0^U + p_1 + p_1 p_0 L_0^U + p_1^2 (1 + p_0 L_0^U + p_1 L_3^U)$$
$$L_0^U = 1 + p_0 L_0^U + p_1 + p_1 p_0 L_0^U + p_1^2 + p_1^2 p_0 L_0^U + p_1^3 L_3^U$$
$$\vdots$$
$$L_0^U = 1 + p_0 L_0^U + p_1 + p_1 p_0 L_0^U + p_1^2 + p_1^2 p_0 L_0^U + p_1^3 + ... + p_1^{r-1} L_{r-1}^U$$
$$L_0^U = 1 + p_0 L_0^U + p_1 + p_1 p_0 L_0^U + p_1^2 + p_1^2 p_0 L_0^U + p_1^3 + ... + p_1^{r-1} (1 + p_0 L_0^U)$$
$$L_0^U = 1 + p_0 L_0^U + p_1 + p_1 p_0 L_0^U + p_1^2 + p_1^2 p_0 L_0^U + p_1^3 + ... + p_1^{r-1} + p_1^{r-1} p_0 L_0^U$$

$$L_0^U - p_0 L_0^U - p_1 p_0 L_0^U - p_1^2 p_0 L_0^U - ... - p_1^{r-1} p_0 L_0^U = 1 + p_1 + p_1^2 + p_1^3 + ... + p_1^{r-1} \quad (2.22)$$

It can be proven by induction on r that $\sum_{i=0}^{r-1} p_1^{\,i} = \frac{1-p_1^{\,r}}{1-p_1}$ for $p_1 \neq 1$. Making use of this, expression (2.22) can be simplified to

$$L_0^U - p_0 L_0^U \left(\frac{1-p_1^{\,r}}{1-p_1} \right) = \frac{1-p_1^{\,r}}{1-p_1}$$
$$L_0^U \left[\frac{1-p_1-p_0(1-p_1^{\,r})}{1-p_1} \right] = \frac{1-p_1^{\,r}}{1-p_1}$$

$$L_0^U = \frac{1-p_1^{\,r}}{1-p_1-p_0(1-p_1^{\,r})}. \quad (2.23)$$

Expression (2.23) of this thesis is given in Amin et al. (1995) and determined in Page (1962). Expression (2.23) is a closed form expression of the in-control average run length of a one-sided chart with warning and action limits in the positive direction only.

Therefore, the in-control average run length of the one-sided (upper or positive direction) chart with warning limit w_U and control limit (action limit) at a_U is given by (2.23) where

$$p_0 = P(SN_i < w_U) = \sum_{i=0}^{\frac{w+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i} \tag{2.24}$$

and

$$p_1 = P(w_U \le SN_i < a_U) = \sum_{i=0}^{\frac{a+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i} - \sum_{i=0}^{\frac{w+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i}. \tag{2.25}$$

The derivations of (2.24) and (2.25) are given below.

Derivation of expression (2.24):

p_0
$= P(SN_i < w_U)$

by using the relationship between SN_i and T_i (recall that $SN_i = 2T_i - n$) we obtain

$= P(2T_i - n < w_U)$
$= P\left(T_i < \frac{w_U + n}{2}\right)$
$= P\left(T_i \le \frac{w_U + n - 2}{2}\right)$

given that T_i is binomially distributed with parameters n and $p = P(X_{ij} \ge \theta_0)$ we obtain

$= \sum_{i=0}^{\frac{w_U+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i}$

since the upper warning limit is equal to some constant w, i.e. $w_U = w$ (see Section 2.2.1) we obtain

$= \sum_{i=0}^{\frac{w+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i}$.

Derivation of expression (2.25):

p_1

$= P(w_U \le SN_i < a_U)$

by using the relationship between SN_i and T_i (recall that $SN_i = 2T_i - n$) we obtain

$= P(w_U \le 2T_i - n < a_U)$

$$= P\left(\frac{w_U + n}{2} \le T_i < \frac{a_U + n}{2}\right)$$

$$= P\left(T_i \le \frac{a_U + n - 2}{2}\right) - P\left(T_i \le \frac{w_U + n - 2}{2}\right)$$

given that T_i is binomially distributed with parameters n and $p = P(X_{ij} \ge \theta_0)$ we obtain

$$= \sum_{i=0}^{\frac{a_U+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i} - \sum_{i=0}^{\frac{w_U+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i}$$

since the upper action and warning limits are equal to constant values represented by a and w, respectively, i.e. $a_U = a$ and $w_U = w$ (see Section 2.2.1), we obtain

$$= \sum_{i=0}^{\frac{a+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i} - \sum_{i=0}^{\frac{w+n-2}{2}} \binom{n}{i} p^i (1-p)^{n-i} .$$

The in-control average run length of the one-sided (lower or negative direction) chart with warning limit at w_L and control limit (action limit) at a_L can be found by replacing p_0 and p_1 by q_0 and q_1 where $q_0 = P(SN_i > w_L)$ and $q_1 = P(a_L < SN_i \le w_L)$. The in-control average run length for the two-sided chart, denoted L_0, can then be obtained using a result in Roberts (1958), $\frac{1}{L_0} = \frac{1}{L_0^U} + \frac{1}{L_0^L}$. The lower one-sided and two-sided charts with warning limits are discussed in detail in Sections 2.2.3.2 and 2.2.4 respectively.

The in-control average run length for the upper one-sided control chart with both warning and action limits is calculated for a specific example ($n = 10$, $p = 0.5$) by evaluating expressions (2.23), (2.24) and (2.25). These values are shown in Table 2.10. Amin, Reynolds and Bakir

(1995) studied Shewhart charts with both warning and action limits. They constructed a table containing the values of L_0^U for Shewhart charts using the sign statistic when $n = 10$ and $p = 0.5$. The values of L_0^U were noted for $a_U = 8$ and 10, $w_U = 0(2)8$ and $r = 2,3,...,7$. Table 2.10 is similar to the table constructed by Amin, Reynolds and Bakir (1995). Note that the values of L_0^U can also be constructed for other values of n, a_U, w_U and r.

Table 2.10. Values of L_0^U for Shewhart charts with both warning and action limits when $n = 10$ and $p = 0.5$.*

	$a_U = 8$				$a_U = 10$			
w_U / r	**0**	**2**	**4**	**6**	**2**	**4**	**6**	**8**
2	4.1	9.2	30.2	79.4	9.6	38.6	269.2	933.7
3	7.9	23.0	70.1	92.4	27.8	194.7	890.3	1023.0
4	13.5	44.7	88.4	93.1	73.0	593.7	1015.8	1024.0
5	21.2	66.9	92.3	93.1	175.4	911.2	1023.6	1024.0
6	31.0	81.5	93.0	93.1	364.4	1002.8	1024.0	1024.0
7	42.2	88.5	93.1	93.1	609.8	1020.3	1024.0	1024.0

Studying Table 2.10 we observe the following. For values of w_U close to a_U and r reasonably large, the introduction of warning lines will have little effect on L_0^U. The reason being that if w_U is close to a_U, the probability of having r successive points plot in this small interval $[w_u, a_U)$ is small. As an example, the calculation of the in-control average run length for $n = 10$, $p = 0.5$, $a_U = 8$, $w_U = 2$ and $r = 6$ will be given. By substituting these values into equations (2.23), (2.24) and (2.25) we obtain

$$p_0 = \sum_{i=0}^{5} \binom{10}{i} (0.5)^{10} = 0.6230,$$

$$p_1 = \sum_{i=0}^{8} \binom{10}{i} (0.5)^{10} - \sum_{i=0}^{5} \binom{10}{i} (0.5)^{10} = 0.3662 \text{ and}$$

$$L_0^U = \frac{1 - p_1^{\,r}}{1 - p_1 - p_0(1 - p_1^{\,r})} = \frac{1 - (0.3662)^6}{1 - 0.3662 - 0.6230\left(1 - (0.3662)^6\right)} = 81.4689 \approx 81.5.$$

* See Mathcad Program 3 in Appendix B for the calculation of the values in Table 2.10. This table also appears in Amin, Reynolds and Bakir (1995), page 1606, Table 2.

2.2.3.2. Lower one-sided control charts

The lower one-sided Shewhart sign chart, described previously, is efficient in detecting large process shifts quickly. Since it is known to be inefficient in detecting small process shifts, a lower warning limit is drawn above the lower action limit to increase its sensitivity for detecting small shifts.

Define rule A_L as: 'Action will be taken if r successive points fall between a_L and w_L (denoted by $R(r, r, a_L, w_L)$) *or* if any point falls below a_L (denoted by $R(1,1,-\infty, a_L)$)'. Clearly, rule A_L is created to detect downward shifts. Let L^L denote the ARL of rule A_L. L_0^L can be computed similarly as L_0^U (see equation (2.23)) with p_0 and p_1 being replaced by q_0 and q_1, where q_0 denotes the probability that a given sample point falls above w_L and q_1 denotes the probability that a given sample point falls between a_L and w_L. Therefore, we have that

$$L_0^L = \frac{1-{q_1}^r}{1-q_1-q_0(1-{q_1}^r)} \tag{2.26}$$

where

$$q_0 = P(SN_i > w_L) = 1 - \sum_{i=0}^{\frac{n-w}{2}} \binom{n}{i} p^i (1-p)^{n-i} \tag{2.27}$$

and

$$q_1 = P(a_L < SN_i \le w_L) = \sum_{i=0}^{\frac{n-w}{2}} \binom{n}{i} p^i (1-p)^{n-i} - \sum_{i=0}^{\frac{n-a}{2}} \binom{n}{i} p^i (1-p)^{n-i}. \tag{2.28}$$

Compare expressions (2.27) and (2.28) to (2.24) and (2.25).

The in-control average run length for the lower one-sided control chart with both warning and action limits is calculated for a specific example ($n = 10$, $p = 0.5$) by evaluating expressions (2.26), (2.27) and (2.28). These values are shown in Table 2.11.

Table 2.11. Values of L_0^L for Shewhart charts with both warning and action limits when $n=10$ and $p=0.5$.*

	$a_L=8$				$a_L=10$			
w_L / r	**0**	**2**	**4**	**6**	**2**	**4**	**6**	**8**
2	4.1	9.2	30.2	79.4	9.6	38.6	269.2	933.7
3	7.9	23.0	70.1	92.4	27.8	194.7	890.3	1023.0
4	13.5	44.7	88.4	93.1	73.0	593.7	1015.8	1024.0
5	21.2	66.9	92.3	93.1	175.4	911.2	1023.6	1024.0
6	31.0	81.5	93.0	93.1	364.4	1002.8	1024.0	1024.0
7	42.2	88.5	93.1	93.1	609.8	1020.3	1024.0	1024.0

Studying Table 2.11 we observe the following. For values of w_L close to a_L and r reasonably large, the introduction of warning lines will have little effect on L_0^L. The reason being that if w_L is close to a_L, the probability of having r successive points plot in this small interval $(a_L, w_L]$ is small. As stated earlier, due to the symmetry of the Binomial distribution we have that if $a_U = a$ then let $a_L = -a$ and if $w_U = w$ then let $w_L = -w$. As a result the values of L_0^L and the values of L_0^U are equal.

As an example, the calculation of the in-control average run length for $n=10$, $p=0.5$, $a_L=8$, $w_L=2$ and $r=6$ will be given. By substituting these values into equations (2.26), (2.27) and (2.28) we obtain

$$q_0 = 1-\sum_{i=0}^{4}\binom{10}{i}(0.5)^{10} = 0.6230,$$

$$q_1 = \sum_{i=0}^{4}\binom{10}{i}(0.5)^{10} - \sum_{i=0}^{1}\binom{10}{i}(0.5)^{10} = 0.3662 \text{ and}$$

$$L_0^L = \frac{1-q_1^{\,r}}{1-q_1-q_0(1-q_1^{\,r})} = \frac{1-(0.3662)^6}{1-0.3662-0.6230\left(1-(0.3662)^6\right)} = 81.4689 \approx 81.5.$$

* See Mathcad Program 3 in Appendix B for the calculation of the values in Table 2.11.

2.2.4. Two-sided control charts

Roberts (1958) provided a method of approximating the *ARL* of the two-sided Shewhart chart with both warning and action limits. The *ARL* for each separate one-sided Shewhart chart was calculated and then combined by applying equation (2.29)

$$\frac{1}{L_0} = \frac{1}{L_0^U} + \frac{1}{L_0^L}. \tag{2.29}$$

(See Appendix A Theorem 1 for a step-by-step derivation of equation (2.29)). Equation (2.29) can be re-written as

$$L_0 = \frac{L_0^L L_0^U}{L_0^L + L_0^U} \tag{2.30}$$

where L_0 denotes the *ARL* of a two-sided chart. In practice some observations can be tied with the specified median. If the number of such cases, within a sample, is small (relative to *n*) one can drop the tied cases and reduce *n* accordingly. On the other hand, if the number of ties is large, more sophisticated analysis might be necessary.

The in-control average run length for the two-sided control chart with both warning and action limits is calculated by evaluating expression (2.30). These values are shown in Table 2.12 for $n = 10$, $p = 0.5$, $a = 8$ and $10, w = 0(2)8$ and $r = 2,3,...,7$.

Table 2.12. Values of L_0 for Shewhart charts with both warning and action limits when $n = 10$ and $p = 0.5$.*

	$a = 8$				$a = 10$			
w / ***r***	**0**	**2**	**4**	**6**	**2**	**4**	**6**	**8**
2	2.1	4.6	15.1	39.7	4.8	19.3	134.6	466.9
3	4.0	11.5	35.0	46.2	13.9	97.4	445.2	511.5
4	6.7	22.4	44.2	46.5	36.5	296.8	507.9	512.0
5	10.6	33.5	46.2	46.5	87.7	455.6	511.8	512.0
6	15.5	40.7	46.5	46.5	182.2	501.4	512.0	512.0
7	21.1	44.2	46.5	46.5	304.9	510.2	512.0	512.0

* See Mathcad Program 3 in Appendix B for the calculation of the values in Table 2.12.

Studying Table 2.12 we observe the following: For values of w close to a and r reasonably large, the introduction of warning limits will have little effect on L_0. The reason being that if w is close to a (and, consequently, $-w$ is close to $-a$), the probability of having r successive points plot in the small interval $[w,a)$ or $(-a,-w]$ is small. These procedures can't be meaningfully illustrated using the data of Montgomery (2001) because the sample size $n = 5$ used there is too small. It may be noted that the highest possible L_0 values for the basic (without warning limits) two-sided sign chart can be seen to be 2^{n-1} (see Amin, Reynolds and Bakir (1995) pp. 1609-1610 and their Appendix on pp. 1620-1621 for a detailed discussion on and a proof that $max\, L_0 = 2^{n-1}$). Thus, achievable values of L_0 are too small for practical use, unless n is about 10. In Table 2.13, the charting constants, i.e. the warning and action limits, are shown, along with the achieved *ARL* values, for the in-control and one out-of-control case. The *ARL* values for the two-sided sign chart, without the warning limits, are shown in each case, within parentheses, for reference.

Table 2.13. In-control *ARL* values for the two-sided sign chart with and without warning limits for $n = 10$*.

	$r = 2$	**$r = 3$**	**$r = 6$**
	$w = 7$ and $a = 10$		
$p = 0.5$ (in-control)	208.97 (512.00)	476.03 (512.00)	511.99 (512.00)
$p = 0.6$ (out-of-control)	35.03 (162.60)	103.28 (162.60)	162.17 (162.60)
	$w = 7$ and $a = 8$		
$p = 0.5$ (in-control)	42.86 (46.55)	46.37 (46.55)	46.55 (46.55)
$p = 0.6$ (out-of-control)	16.37 (20.82)	20.16 (20.82)	20.82 (20.82)

It is seen that adding warning limits to a control chart decreases its average run length. For example, adding a warning limit at 7 to the basic sign chart with an action limit at 10 decreases the ARL_0 approximately 59% (from 512 to 208.97), when $r = 2$, 7% (512 to 476.03) when $r = 3$ and 0.002% (512 to 511.99) when $r = 6$, respectively. The out-of-control average

* See Mathcad Program 3 in Appendix B for the calculation of the values in Table 2.13.

run length is decreased by approximately 79% (from 162.6 to 35.03) when $r = 2$, 36% (from 162.6 to 103.28) when $r = 3$ and 0.26% (from 162.6 to 162.17) when $r = 6$, respectively. Note that although the out-of-control average run length is reduced significantly (which means a quicker detection of shift) by the addition of warning limits, the ARL_0 is also reduced significantly. This poses a dilemma in practice, since it is desirable to have a high ARL_0 and a low *FAR*, so one would need to strike a balance. One possibility is to use warning limits closer to the action limits. For example, from the second panel of Table 2.13, we see that adding a warning limit at 7 to the sign chart with an action limit of 8, decreases the ARL_0 by only 8% (from 46.55 to 42.86) when $r = 2$ and has little effect on ARL_0 when r is reasonably large. Amin et al. (1995) concluded that for the upper one-sided Shewhart-type sign chart, introduction of warning limits will have little effect on the in-control average run length, but can significantly reduce the out-of-control average run length for small shifts when the warning limits are chosen close to the action limits and r is reasonably large. Similar conclusions are expected to hold for two-sided charts.

Up to this point we have discussed methods to increase the sensitivity of standard Shewhart control charts to small process shifts. Another method is to extend the existing charts by incorporating various signaling rules involving runs of the plotting statistic. The signaling rules considered include the following: A process is declared to be out-of-control when (a) a single point (charting statistic) plots outside the control limit(s) (*1-of-1* rule) (b) k consecutive points (charting statistics) plot outside the control limit(s) (*k-of-k* rule) or (c) exactly k of the last w points (charting statistics) plot outside the control limit(s) (*k-of-w* rule). We can consider these signaling rules where both k and w are positive integers with $1 \le k \le w$ and $w \ge 2$. Rule (a) is the simplest and is the most frequently used in the literature. Thus, the *1-of-1* rule corresponds to the usual control chart, where a signal is given when a plotting statistic falls outside the control limit(s). Rules (a) and (b) are special cases of rule (c); rules (b) and (c) have been used in the context of supplementing the Shewhart charts with warning limits and zones.

Example 2.4

A two-sided Shewhart chart incorporating the *2-of-2* rule with *one* absorbing state that corresponds to the out-of-control signal

In this example, a control chart is viewed as consisting of the zones shown in Figure 2.4.

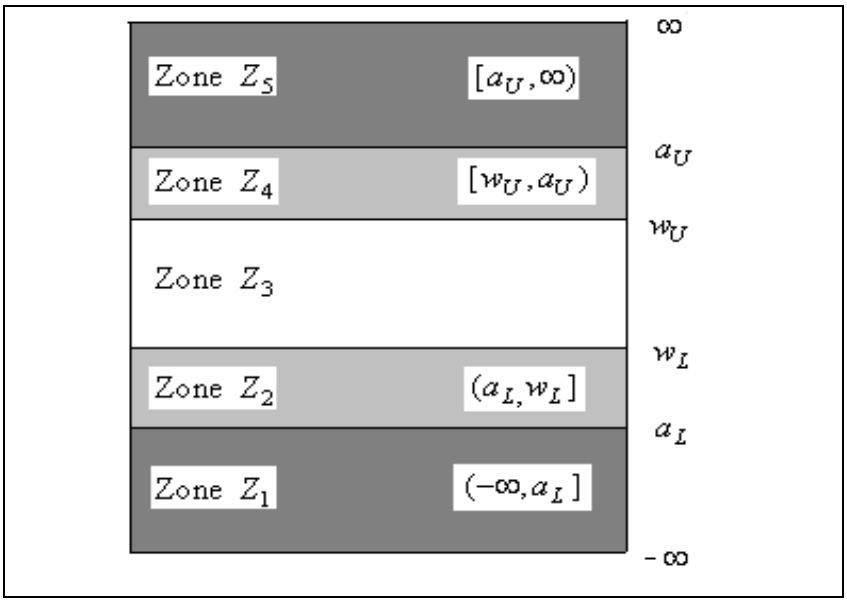

Figure 2.4. A control chart partitioned into 5 zones.

Let p_i denote the probability of plotting in Zone Z_i for $i = 1,2,3,4,5$. To illustrate the calculation of signal probabilities, the following set of rules is used:

$$\{R(1,1,-\infty,a_L) \cup R(2,2,a_L,w_L) \cup R(2,2,w_U,a_U) \cup R(1,1,a_U,\infty)\}.$$

$R(1,1,-\infty,a_L)$: The chart will signal if any 1 point falls in Zone Z_1 (below LCL).

$R(2,2,a_L,w_L)$: The chart will signal if any 2 successive points fall in Zone Z_2.

$R(2,2,w_U,a_U)$: The chart will signal if any 2 successive points fall in Zone Z_4.

$R(1,1,a_U,\infty)$: The chart will signal if any 1 point falls in Zone Z_5 (above UCL).

Table 2.14. Classifications and descriptions of states.

State number	Description of state	Absorbent (A)/ Non-absorbent (NA)
0	No points beyond any of the control limits. Point plots in Zone Z_3	NA
1	Point plots in Zone Z_2	NA
2	Point plots in Zone Z_4	NA
3	Point plots below a_L or above a_U or 2 successive points fall between w_U and a_U or 2 successive points fall between w_L and a_L.	A

Clearly, $\sum_{i=1}^{5} p_i = p_1 + p_2 + p_3 + p_4 + p_5 = 1$, since the statistic **must** plot in one of the 5 zones. The transition probabilities are given in the transition probability matrix, $TPM = [p_{ij}]$, for $i = 0,1,2,3$ and $j = 0,1,2,3$.

$$TPM_{4\times4} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} p_3 & p_2 & p_4 & p_1 + p_5 \\ p_3 & 0 & p_4 & p_1 + p_5 + p_2 \\ p_3 & p_2 & 0 & p_1 + p_5 + p_4 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

From (2.18) we have $\sum_{j\in\Omega} p_{ij} = 1 \quad \forall i$. This is easily shown for $i = 0,1,2,3$. For example, for $i = 0$ we have that $\sum_{j=0}^{3} p_{0j} = p_3 + p_2 + p_4 + p_1 + p_5 = 1$. The rest of the calculations follow similarly. Table 2.15 illustrates that the TPM can be partitioned into 4 sections.

Table 2.15. Transition probabilities of the *2-of-2* rule for a Markov chain with *one* absorbing state.

	States at time $t+1$			
States at time t	0 (NA)	1 (NA)	2 (NA)	3 (A)
0 (NA)	p_3	p_2	p_4	p_1+p_5
1 (NA)	p_3	0	p_4	$p_1+p_5+p_2$
2 (NA)	p_3	p_2	0	$p_1+p_5+p_4$
3 (A)	0	0	0	1

=

States at time $t+1$			
0 (NA)	1 (NA)	2 (NA)	3 (A)
$Q_{3\times3}$			$\underline{p}_{3\times1}$
$\underline{0}'_{1\times3}$			$1_{1\times1}$

Brook and Evans (1972) showed that the *ARL* for initial state i can be calculated by adding the elements in the i^{th} row of $(I_{3\times3}-Q_{3\times3})^{-1}$. Making use of $(I-Q)^{-1}$ is typically done in stochastic processes where one works with recurrence and first passage times (see, for example, Bartlett (1953)).

$$I_{3\times3}-Q_{3\times3}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}p_3&p_2&p_4\\p_3&0&p_4\\p_3&p_2&0\end{pmatrix}=\begin{pmatrix}1-p_3&-p_2&-p_4\\-p_3&1&-p_4\\-p_3&-p_2&1\end{pmatrix}$$

$$(I_{3\times3}-Q_{3\times3})^{-1}=\begin{pmatrix}1-p_3&-p_2&-p_4\\-p_3&1&-p_4\\-p_3&-p_2&1\end{pmatrix}^{-1}=$$

$$\frac{-1+p_4p_2}{(-1+p_4p_2+p_3+p_3p_4p_2+p_3p_2+p_3p_4)}\begin{pmatrix}-1+p_4p_2&-p_2(1+p_4)&-p_4(1+p_2)\\-p_3(1+p_4)&-1+p_3+p_3p_4&-p_4\\-p_3(1+p_2)&-p_2&-1+p_3+p_3p_2\end{pmatrix}.$$

The *ARL* for initial state i can be calculated by adding the elements in the i^{th} row of $(I_{3\times3}-Q_{3\times3})^{-1}$ for $i=1,2,3$.

Example 2.5

A two-sided Shewhart chart incorporating the *2-of-2* rule with *more than one* absorbing state that corresponds to an out-of-control signal.

This example is similar to the previous example in having three transient states, but differs from the previous example by having *more than one* absorbing state. By changing the classification of the states, the generalization to more than one absorbing state is considered. We need to introduce a rule number and this is done by adding a subscript to each rule, i.e. in general we have that $R(r,n,k,l)$ which now becomes $R_j(r,n,k,l)$ where *j* denotes the rule number. This modification allows for a separate absorbing state, A_j, that is associated with each of the runs rules, $R_j(r,n,k,l)$. This modification of Champ and Woodall's (1987) method was done by Champ and Woodall (1997). As a result, we have the following rules with the corresponding absorbing states (see Figure 2.4 for the partitioning of the control chart into 5 zones):

- $R_1(1,1,-\infty,a_L)$ associated with absorbing state A_1: The chart will signal if any 1 point falls in Zone Z_1 (below LCL).
- $R_2(2,2,a_L,w_L)$ associated with absorbing state A_2: The chart will signal if any 2 successive points fall in Zone Z_2.
- $R_3(2,2,w_U,a_U)$ associated with absorbing state A_3: The chart will signal if any 2 successive points fall in Zone Z_4.
- $R_4(1,1,a_U,\infty)$ associated with absorbing state A_4: The chart will signal if any 1 point falls in Zone Z_5 (above UCL).

Table 2.16. States and next-state transitions by zones.

		Zones					
State number	**State vector**	**Z_1** $(-\infty, a_L]$	**Z_2** $(a_L, w_L]$	**Z_3** (w_L, w_U)	**Z_4** $[w_U, a_U)$	**Z_5** $[a_U, \infty)$	**Absorbent (A)/ Non-absorbent (NA)**
0	(0,0)	A_1	(1,0)	(0,0)	(0,1)	A_4	NA
1	(1,0)	A_1	A_2	(0,0)	(0,1)	A_4	NA
2	(0,1)	A_1	(1,0)	(0,0)	A_3	A_4	NA
3	A_1	A_1	A_1	A_1	A_1	A_1	A
4	A_2	A_2	A_2	A_2	A_2	A_2	A
5	A_3	A_3	A_3	A_3	A_3	A_3	A
6	A_4	A_4	A_4	A_4	A_4	A_4	A

Each non-absorbing state in Table 2.16 is represented by a vector of 0's and 1's. The vector indicates by the 1's only those observations that may contribute to an out-of-control signal. Let p_i denote the probability of plotting in Zone Z_i for $i = 1,2,3,4,5$. Clearly, $\sum_{i=1}^{5} p_i = p_1 + p_2 + p_3 + p_4 + p_5 = 1$, since the statistic **must** plot in one of the 5 zones. The transition probabilities are given in the transition probability matrix, $TPM = [p_{ij}]$, for $i = 0,1,...,6$ and $j = 0,1,...,6$.

$$TPM_{7\times 7} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} & p_{05} & p_{06} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} & p_{15} & p_{16} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & p_{26} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & p_{36} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} \\ p_{50} & p_{51} & p_{52} & p_{53} & p_{54} & p_{55} & p_{56} \\ p_{60} & p_{61} & p_{62} & p_{63} & p_{64} & p_{65} & p_{66} \end{pmatrix} = \begin{pmatrix} p_3 & p_2 & p_4 & p_1 & 0 & 0 & p_5 \\ p_3 & 0 & p_4 & p_1 & p_2 & 0 & p_5 \\ p_3 & p_2 & 0 & p_1 & 0 & p_4 & p_5 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

From (2.18) we have $\sum_{j\in\Omega} p_{ij} = 1 \quad \forall i$. This is easily shown for $i = 0,1,...,6$. For example, for $i = 0$ we have that $\sum_{j=0}^{6} p_{0j} = p_3 + p_2 + p_4 + p_1 + 0 + 0 + p_5 = 1$. The rest of the calculations follow similarly. Table 2.17 illustrates that the TPM can be partitioned into 4 sections.

Table 2.17. Transition probabilities of the *2-of-2* rule for a Markov chain with more than one absorbing state.

	States at time $t+1$						
States at time t	0 (NA)	1 (NA)	2 (NA)	3 (A)	4 (A)	5 (A)	6 (A)
0 (NA)	p_3	p_2	p_4	p_1	0	0	p_5
1 (NA)	p_3	0	p_4	p_1	p_2	0	p_5
2 (NA)	p_3	p_2	0	p_1	0	p_4	p_5
3 (A)	0	0	0	1	0	0	0
4 (A)	0	0	0	0	1	0	0
5 (A)	0	0	0	0	0	1	0
6 (A)	0	0	0	0	0	0	1

=

States at time $t+1$						
0 (NA)	1 (NA)	2 (NA)	3 (A)	4 (A)	5 (A)	6 (A)
$Q_{3\times3}$			$C_{3\times4}$			
$Z_{4\times3}$			$I_{4\times4}$			

where the essential transition probability sub-matrix $Q_{3\times3} = \begin{pmatrix} p_3 & p_2 & p_4 \\ p_3 & 0 & p_4 \\ p_3 & p_2 & 0 \end{pmatrix}$ contains all the transition probabilities of going from a non-absorbent state to a non-absorbent state, $Q:(NA \to NA)$. $C_{3\times4} = \begin{pmatrix} p_1 & 0 & 0 & p_5 \\ p_1 & p_2 & 0 & p_5 \\ p_1 & 0 & p_4 & p_5 \end{pmatrix}$ contains all the transition probabilities of going from each non-absorbent state to the absorbent states, $C:(NA \to A)$. $Z_{4\times3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ contains all the transition probabilities of going from each absorbent state to the non-absorbent states, $Z:(A \to NA)$. The Z matrix is the zero matrix, since it is impossible to go from an absorbent state to a non-absorbent state, because once an absorbent state is entered, it is never left. $I_{4\times4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ contains all the transition probabilities of going from an absorbent state to an absorbent state, $I:(A \to A)$. A square matrix of this form is called the identity matrix.

$TPM = TPM^1$ represents the probability that the process will, when in state i, next make a transition to state j, in **one** step. Consider the matrix, TPM^n (read: "TPM to the power of n"), the non-absorbing state i and the absorbing state j. The j^{th} component of TPM^n is the probability that a signal will be caused by the j^{th} set of runs rules that can cause a signal on or before the n^{th} sampling stage given the chart begins in non-absorbing state i. For that reason, an equation for TPM^n is desired. In addition, we will show that $\lim_{n\to\infty} TPM^n = \begin{pmatrix} Z & B \\ Z & I \end{pmatrix}$ where the elements of the matrix B are the probabilities that the chart will go from a non-absorbent state (where no signal is given) to an absorbent state (where a signal is given) in n transitions. We are interested in the matrix B, because the $(i,j)^{th}$ element of B is the long run proportion of times the j^{th} set of runs rules causes the chart to signal given the chart starts in a non-absorbing state i.

Probabilities on the n^{th} sampling stage

$$TPM_{7\times7} = TPM^1_{7\times7} = \begin{pmatrix} Q_{3\times3} & C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}$$

$$\begin{aligned} & TPM^2_{7\times7} \\ &= TPM^1_{7\times7} \cdot TPM^1_{7\times7} \\ &= \begin{pmatrix} Q_{3\times3} & C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}\begin{pmatrix} Q_{3\times3} & C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix} \\ &= \begin{pmatrix} Q^2_{3\times3} + C_{3\times4}Z_{4\times3} & Q_{3\times3}C_{3\times4} + C_{3\times4}I_{4\times4} \\ Z_{4\times3}Q_{3\times3} + I_{4\times4}Z_{4\times3} & Z_{4\times3}C_{3\times4} + I^2_{4\times4} \end{pmatrix} = \begin{pmatrix} Q^2_{3\times3} & Q_{3\times3}C_{3\times4} + C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
&TPM^{\,3}_{7\times7}\\
&= TPM^{\,2}_{7\times7}\cdot TPM^{\,1}_{7\times7}\\
&= \begin{pmatrix} Q^2_{3\times3} & Q_{3\times3}C_{3\times4}+C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}\begin{pmatrix} Q_{3\times3} & C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}\\
&= \begin{pmatrix} Q^3_{3\times3}+Q_{3\times3}C_{3\times4}Z_{4\times3}+C_{3\times4}Z_{4\times3} & Q^2_{3\times3}C_{3\times4}+Q_{3\times3}C_{3\times4}I_{4\times4}+C_{3\times4}I_{4\times4} \\ Z_{4\times3}Q_{3\times3}+I_{4\times4}Z_{4\times3} & Z_{4\times3}C_{3\times4}+I^2_{4\times4} \end{pmatrix}\\
&= \begin{pmatrix} Q^3_{3\times3} & Q^2_{3\times3}C_{3\times4}+Q_{3\times3}C_{3\times4}+C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}
\end{aligned}
$$

continuing in this way, we obtain

$$TPM^{\,n}_{7\times7} = \begin{pmatrix} Q^n_{3\times3} & Q^{n-1}_{3\times3}C_{3\times4}+Q^{n-2}_{3\times3}C_{3\times4}+\ldots+Q_{3\times3}C_{3\times4}+C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix} = \begin{pmatrix} Q^n_{3\times3} & (Q^{n-1}_{3\times3}+Q^{n-2}_{3\times3}+\ldots+Q_{3\times3}+I_{3\times3})C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}.$$

This expression can be simplified by applying the following Corollary.

Corollary 2.1

$$(Q^{n-1}+Q^{n-2}+\ldots Q+I)C = (I-Q)^{-1}(I-Q^n)C$$

Proof:

$$
\begin{aligned}
&(I-Q)(I+Q+Q^2+Q^3+\ldots+Q^{n-1})\\
&= I+Q+Q^2+Q^3+\ldots+Q^{n-1}-(Q+Q^2+Q^3+\ldots+Q^{n-1}+Q^n)\\
&= I+Q+Q^2+Q^3+\ldots+Q^{n-1}-Q-Q^2-Q^3-\ldots-Q^{n-1}-Q^n\\
&= I-Q^n
\end{aligned}
$$

$$TPM^{\,n}_{7\times7} = \begin{pmatrix} Q^n_{3\times3} & (Q^{n-1}_{3\times3}+Q^{n-2}_{3\times3}+\ldots+Q_{3\times3}+I_{3\times3})C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix} = \begin{pmatrix} Q^n_{3\times3} & (I_{3\times3}-Q_{3\times3})^{-1}(I_{3\times3}-Q^n_{3\times3})C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}$$

$$\lim_{n\to\infty} TPM^{\,n}_{7\times7} = \begin{pmatrix} \lim_{n\to\infty} Q^n_{3\times3} & \lim_{n\to\infty}(I_{3\times3}-Q_{3\times3})^{-1}(I_{3\times3}-Q^n_{3\times3})C_{3\times4} \\ \lim_{n\to\infty} Z_{4\times3} & \lim_{n\to\infty} I_{4\times4} \end{pmatrix} = \begin{pmatrix} Z_{3\times3} & (I_{3\times3}-Q_{3\times3})^{-1}C_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix} = \begin{pmatrix} Z_{3\times3} & B_{3\times4} \\ Z_{4\times3} & I_{4\times4} \end{pmatrix}$$

where $B_{3\times4} = (I_{3\times3}-Q_{3\times3})^{-1}C_{3\times4}$.

$\lim_{n\to\infty} Q_{3\times3}^{n} = Z_{3\times3}$, because the elements of $Q_{3\times3}^{n}$ are the transition probabilities that the chart will go from a non-absorbent state to a non-absorbent state in n transitions. These probabilities will tend to 0 as n tends to infinity, because once the system has moved from a non-absorbent state to an absorbent state, that absorbent state can't be left, i.e. the system will not be able to move back to a non-absorbent state.

Recall that we are interested in the matrix B, because the $(i,j)^{th}$ element of B is the long run proportion of times the j^{th} set of runs rules causes the chart to signal given the chart starts in a non-absorbing state i.

$$B_{3\times4} = (I_{3\times3} - Q_{3\times3})^{-1} C_{3\times4}$$

$$= \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} p_3 & p_2 & p_4 \\ p_3 & 0 & p_4 \\ p_3 & p_2 & 0 \end{pmatrix} \right]^{-1} \begin{pmatrix} p_1 & 0 & 0 & p_5 \\ p_1 & p_2 & 0 & p_5 \\ p_1 & 0 & p_4 & p_5 \end{pmatrix}$$

$$= \begin{pmatrix} 1-p_3 & -p_2 & -p_4 \\ -p_3 & 1 & -p_4 \\ -p_3 & -p_2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} p_1 & 0 & 0 & p_5 \\ p_1 & p_2 & 0 & p_5 \\ p_1 & 0 & p_4 & p_5 \end{pmatrix}$$

$$= \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix}.$$

The effort in inverting $(I-Q)$ could be substantial and therefore some type of statistical software package is desirable. Using Mathcad's **Symbolics → Evaluate → Symbolically** we can easily calculate the inverse of $(I-Q)$ and multiply the two matrices, $(I-Q)^{-1}$ and C, to get an expression for the matrix B. The long run signal probabilities are given by b_{11}, b_{12}, b_{13}, b_{14}, b_{21}, b_{22}, b_{23}, b_{24}, b_{31}, b_{32}, b_{33} and b_{34}. Since these are all very long expressions, only one will be given and explained: $b_{11} = \dfrac{p_1(1+p_2)(1+p_4)}{1-p_2p_4-p_3-p_2p_3p_4-p_2p_3-p_3p_4}$ is the long run proportion of times that the runs rule R_1 causes the chart to signal when the chart

starts in state 1. In general: b_{ij} is the long run proportion of times that the runs rule R_j causes the chart to signal when the chart starts in state i.

2.2.5. Summary

The necessary steps for calculating the probability that any subset of runs rules will give an out-of -control signal:

STEP 1: Classification of states:

- State number
- Description of state
- Absorbent (A) / Non-absorbent (NA)

STEP 2: Setting up the transition probability matrix $TPM = [p_{ij}]$

STEP 3: Partitioning of the transition probability matrix into 4 sections $TPM = [p_{ij}] = \begin{pmatrix} Q & C \\ Z & I \end{pmatrix}$

- $Q : (NA \rightarrow NA)$
- $C : (NA \rightarrow A)$
- $Z : (A \rightarrow NA)$
- $I : (A \rightarrow A)$

STEP 4: Obtain $(I - Q)^{-1}$

STEP 5: Calculate $B = (I - Q)^{-1} C$

STEP 6: Interpret B. b_{ij} is the long run proportion of times that the runs rule R_j causes the chart to signal when the chart starts in state i.

2.3. The tabular CUSUM control chart

2.3.1. Introduction

Cumulative sum (or CUSUM) control charts were first introduced by Page (1954) (although not in its present form) and have been studied by many authors, for example, Barnard (1959), Ewan and Kemp (1960), Johnson (1961), Goldsmith and Whitfield (1961), Page (1961), Ewan (1963), Van Dobben de Bruyn (1968), Woodall and Adams (1993) and Hawkins and Olwell (1998). Montgomery (2005) related CUSUM ideas to other SPC methodologies.

The statistical design of CUSUM 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. The CUSUM chart incorporates all the information in the sequence of sample values by plotting a function of the cumulative sums of the deviations of the sample values from a target value. For example, suppose that samples of size $n = 1$ are collected and let x_j denote the j^{th} observation. The case of individual observations occurs very often in practice, so that situation will be treated first. Later we will see how to modify these results for subgroups. Then if θ_0 is the target value, the CUSUM chart is formed by plotting C_i where

$$C_i = \sum_{j=1}^{i} (x_j - \theta_0) = (x_i - \theta_0) + \sum_{j=1}^{i-1} (x_j - \theta_0) = (x_i - \theta_0) + C_{i-1}.$$

The upper one-sided CUSUM works by accumulating deviations from $\theta_0 + K$ that are above target. For the upper one-sided CUSUM chart we use

$$C_i^+ = \max[0, C_{i-1}^+ + (x_i - \theta_0) - K] \quad \text{for } i = 1,2,3,... \tag{2.31}$$

to detect positive deviations from θ_0. A signaling event occurs for the first i such that $C_i^+ \geq H$.

The lower one-sided CUSUM works by accumulating deviations from $\theta_0 - K$ that are below target. For the lower one-sided CUSUM chart we use

$$C_i^- = \min[0, C_{i-1}^- + (x_i - \theta_0) + K] \quad \text{for } i = 1,2,3,... \tag{2.32}$$

or

$$C_i^{-*} = \max[0, C_{i-1}^{-*} - (x_i - \theta_0) - K] \quad \text{for } i = 1,2,3,... \tag{2.33}$$

to detect negative deviations from θ_0. A signaling event occurs for the first i such that $C_i^- \leq -H$ (if expression (2.32) is used) or $C_i^{-*} \geq H$ (if expression (2.33) is used). For a visually appealing chart, expression (2.32) will be used to construct the lower one-sided CUSUM.

The two-sided CUSUM chart signals for the first i at which either one of the two inequalities is satisfied, that is, either $C_i^+ \geq H$ or $C_i^- \leq -H$. Both K and H are non-negative integers and they are needed in order to implement the CUSUM chart. Details regarding how to choose these constants are given in Section 2.3.1 in the sub-section called *Recommendations for the design of the CUSUM control chart.*

Note that both C_i^+ and C_i^- accumulate deviations from the target value θ_0 that are greater than K. Originally, Page (1954) set the starting values equal to zero, that is, $C_0^+ = 0$ and $C_0^- = 0$. Later on, Lucas and Crosier (1982) recommended setting the starting values equal to some nonzero value to improve the sensitivity of the CUSUM at process start-up. This is referred to as the fast initial response (FIR) or head start feature.

The standardized CUSUM

The variable x_i can be standardized by subtracting its mean and dividing by its standard deviation, that is,

$$y_i = \frac{(x_i - \theta_0)}{\sigma}. \tag{2.34}$$

The resulting standardized upper one-sided CUSUM is given by

$$S_i^+ = \max[0, S_{i-1}^+ + y_i - k] \quad \text{for } i = 1,2,3,... \tag{2.35}$$

while the resulting standardized lower one-sided CUSUM is given by

$$S_i^- = \min[0, S_{i-1}^- + y_i + k] \quad \text{for } i = 1,2,3,... \tag{2.36}$$

or

$$S_i^{-*} = \max[0, S_{i-1}^{-*} - y_i - k] \quad \text{for } i = 1,2,3,... \tag{2.37}$$

The two-sided standardized CUSUM is constructed by running the upper and lower one-sided standardized CUSUM charts simultaneously and signals at the first i such that $S_i^+ \geq H_s$ or $S_i^- \leq -H_s$*. Both k and H_s are non-negative integers and they are needed in order to implement the standardized CUSUM chart. As mentioned previously, details regarding how to choose these constants are given in Section 2.3.1 in the sub-section called *Recommendations for the design of the CUSUM control chart*.

The unstandardized CUSUM C_i and the standardized CUSUM S_i contains the same information. The question arises: Should unstandardized or standardized data be used? Unstandardized data has the advantage that the units of the vertical axis are in their original measurements which makes interpretation easier. Standardized data has the advantage that different CUSUM charts can be compared.

The CUSUM for monitoring the process mean and other sample statistics

A CUSUM chart for monitoring the process mean can be obtained by replacing x_i in expression (2.34) with the sample average $\bar{x}_i$ and by replacing σ by $\sigma/\sqrt{n}$. It is also possible to develop CUSUM charts for other sample statistics, for example, standard deviations and defects. These CUSUM charts for other sample statistics have been studied by many authors, for example, Lucas (1985), Gan (1993) and White, Keats and Stanley (1997).

Recommendations for the design of the CUSUM control chart

Phase II CUSUM charts should be designed on the basis of *ARL* performance. The parameters K and H are obtained for a specified in-control average run length. Both

* The vertical axis of the standardized CUSUM will be measured in multiples of the standard deviation (σ) of the data, whereas the vertical axis of the unstandardized data will be measured in the same units of *X*, for example, in meters, millimeters, ect. To avoid confusion, *H* and H_s will be used to denote the decision intervals for the unstandardized and standardized CUSUM charts, respectively.

parameters are non-negative integers. Let σ denote the standard deviation of the sample variable used in forming the cumulative sum. Parametric CUSUM charts (see Page (1954)) are used for detecting shifts in a normal mean based on the cumulative sum of differences from target. Let $H = h\sigma$ and $K = k\sigma$ where h is usually taken to be equal to 4 or 5 and k is usually taken to equal 0.5 (see Montgomery (2005) page 395). By choosing $h = 4$ or $h = 5$ and $k = 0.5$ (the values most commonly used in practice) we generally get a good average run length performance for parametric CUSUM charts. In the next section we will show that choosing $h = 4$ or $h = 5$ and $k = 0.5$ is not recommended for nonparametric CUSUM charts, since it usually gives a poor in-control average run length performance. Since we will not be using $H = 4\sigma$ or 5σ and $K = 0.5\sigma$ for nonparametric control charts, we will denote the decision interval and reference value by h and k, respectively, from this point forward.

The proposed nonparametric CUSUM chart

Amin, Reynolds and Bakir (1995) proposed a nonparametric CUSUM chart for the median (or any other percentile) of any continuous population based on sign statistics. Recall that for the i^{th} random sample the plotting statistic in the Shewhart-type chart was $SN_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0)$. The chart proposed by Amin et al. (1995) instead uses the cumulative sum of the statistic SN_i with a stopping rule. They also calculated the *ARL* of the chart using a Markov chain approach where the transition probabilities are calculated via the distribution of the sign statistic, which is of course binomial. The procedure is distribution-free since the in-control distribution of SN_i does not depend on the underlying distribution for all continuous distributions. A CUSUM sign chart can be obtained by replacing y_i in expressions (2.35), (2.36) and (2.37) with SN_i. In other words, for the upper one-sided CUSUM sign chart we use

$$S_i^+ = \max[0, S_{i-1}^+ + SN_i - k] \quad \text{for } i = 1,2,3,... \tag{2.38}$$

to detect positive deviations from the known target value θ_0. A signaling event occurs for the first i such that $S_i^+ \geq h$.

For a lower one-sided CUSUM sign chart we use

$$S_i^- = \min[0, S_{i-1}^- + SN_i + k] \quad \text{for } i = 1,2,3,... \tag{2.39}$$

or

$$S_i^{-*} = \max[0, S_{i-1}^{-*} - SN_i - k] \quad \text{for } i = 1,2,3,... \tag{2.40}$$

to detect negative deviations from the known target value θ_0. A signaling event occurs for the first i such that $S_i^- \leq -h$ (if expression (2.39) is used) or $S_i^{-*} \geq h$ (if expression (2.40) is used).

The corresponding two-sided CUSUM chart signals for the first i at which either one of the two inequalities is satisfied, that is, either $S_i^+ \geq h$ or $S_i^- \leq -h$. Starting values are typically chosen to equal zero, that is, $S_0^+ = S_0^- = 0$.

The constants k and h are obtained for a specified in-control average run length. In-control average run length (ARL_0), standard deviation of the run length ($SDRL_0$), 5^{th}, 25^{th} (the first quartile, Q_1), 50^{th} (the median run length, MRL_0), 75^{th} (the third quartile, Q_3) and 95^{th} percentile values will be computed and tabulated for various values of h and k later on.

2.3.2. One-sided control charts

2.3.2.1. Upper one-sided control charts

Various expressions for the exact run length distribution and its parameters have been given for the normal theory one-sided CUSUM procedure by, for example, Ewan and Kemp (1960), Brook and Evans (1972), Woodall (1983) and Hawkins and Olwell (1998). Many authors have presented various approximations for the run length distribution and its parameters for the one-sided CUSUM procedure. A Markov chain representation of the one-sided CUSUM procedure based on integer-valued cumulative sums is presented in this section. The number of states included in the Markov chain is minimized in order to make the methods as efficient as possible.

Markov chain approach

Brook and Evans (1972) and Amin et al. (1995) considered a method for evaluating the exact average run length and its moments for the upper one-sided CUSUM chart by treating the cumulative sum as a Markov chain with the state space a subset of $\{0,1,2,...,h\}$. Markov techniques have a great advantage as they are adjustable to many runs related problems and they often simplify the solutions to the specific problems they are applied on. Fu, Spiring and Xie (2002) presented three results that must be satisfied before implementing the finite-state Markov chain approach.

Let S_t^+ be a finite-state homogenous Markov chain on the state space Ω^+ with a TPM such that (i) $\Omega^+ = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$ where $0 = \varsigma_0 < \varsigma_1 < ... < \varsigma_{r+s-1} = h$ and ς_{r+s-1} is an absorbent state; (ii) the TPM is given by $TPM = [p_{ij}]$ for $i = 0,1,...,r+s-1$ and $j = 0,1,...,r+s-1$ where r denotes the number of non-absorbent states and s the number of absorbent states, respectively, and (iii) the starting value should be in the "dummy" state with probability one, that is, $P(S_0^+ = \varsigma_0) = 1$, to ensure the process starts in-control . Assume that the Markov chain S_t^+ satisfies conditions (i), (ii) and (iii), then from Fu, Spiring and Xie (2002) and Fu and Lou (2003) we have

$$P(N = n \mid S_0^+ = 0) = \underline{\xi} Q^{n-1}(I - Q)\underline{1} \tag{2.41}$$

$$E(N) = \underline{\xi}(I - Q)^{-1}\underline{1} \tag{2.42}$$

$$E(N^2) = \underline{\xi}(I + Q)(I - Q)^{-2}\underline{1} \tag{2.43}$$

$$\text{var}(N) = E(N^2) - (E(N))^2 = \underline{\xi}(I + Q)(I - Q)^{-2}\underline{1} - \left(\underline{\xi}(I - Q)^{-1}\underline{1}\right)^2 \tag{2.44}$$

$$SDRL = \sqrt{\text{var}(N)} = \sqrt{\underline{\xi}(I + Q)(I - Q)^{-2}\underline{1} - \left(\underline{\xi}(I - Q)^{-1}\underline{1}\right)^2} \tag{2.45}$$

where the essential transition probability sub-matrix Q is the $r \times r$ matrix that contains all the transition probabilities of going from a non-absorbent state to a non-absorbent state, I is the $r \times r$ identity matrix, $\underline{\xi}$ is a $1 \times r$ row vector with 1 at the 1^{st} element and zero elsewhere and $\underline{1}$ is an $r \times 1$ column vector with all elements equal to unity. See Theorem 2 in Appendix A for the derivations done by Fu, Spiring and Xie (2002) and Fu and Lou (2003).

The time that the procedure signals is the first time such that the finite-state Markov chain S_t^+ enters one of the absorbent states where the state space is given by $\Omega^+ = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$, $S_0^+ = 0$ and

$$S_t^+ = \min\{h, \max\{0, S_{t-1}^+ + SN_t - k\}\}. \qquad (2.46)$$

The state corresponding to a signal by the CUSUM chart is called an absorbent state. Clearly, there is only one absorbent state, since the chart signals when S_t^+ falls on or above *h*, i.e. $s = 1$.

The distribution of SN_t can easily be obtained from the binomial distribution (recall that $SN_i = 2T_i - n \quad \forall i$, where T_i is binomially distributed with parameters n and $p = P(X_{ij} \geq \theta_0))$. The binomial probabilities are given in Table G of Gibbons and Chakraborti (2003) and can easily be calculated using some type of statistical software package, for example, Excel or SAS.

Example 2.6

An upper one-sided CUSUM sign chart where the sample size is odd ($n = 5$)

The statistical properties of an upper one-sided CUSUM sign chart with a decision interval of 4 ($h = 4$), a reference value of 1 ($k = 1$) and a sample size of 5 ($n = 5$) is examined. For n odd, the reference value is taken to be odd, because this leads to the sum $\sum(SN_i - k)$ being equal to even values which reduces the size of the state space for the Markov chain. This will halve the size of the matrices of transition probabilities. For $h = 4$ we have that the state space is $\Omega^+ = \{\varsigma_0, \varsigma_1, \varsigma_2\} = \{0,2,4\}$ with $0 = \varsigma_0 < \varsigma_1 < \varsigma_2 = h$. The state space is calculated using equation (2.46) and the calculations are shown in Table 2.18.

Table 2.18. Calculation of the state space when $h=4$, $k=1$ and $n=5$.

SN_t	$S_{t-1}^+ + SN_t - k$	$\max\{0, S_{t-1}^+ + SN_t - k\}$	$S_t^+ = \min\{h, \max\{0, S_{t-1}^+ + SN_t - k\}\}$
-5	-6*	0	0
-3	-4	0	0
-1	-2	0	0
1	0	0	0
3	2	2	2
5	4	4	4

Table 2.19. Classifications and descriptions of the states.

State number	Description of the state	Absorbent (A)/ Non-absorbent (NA)
0	$S_t^+ = 0$	NA
1	$S_t^+ = 2$	NA
2	$S_t^+ = 4$	A

From Table 2.19 we see that there are two non-absorbent states, i.e. $r=2$, and one absorbent state, i.e. $s=1$. Therefore, the corresponding TPM will be a $(r+s)\times(r+s)=3\times 3$ matrix. It can be shown (see Table 2.20) that the TPM is given by

$$TPM_{3\times 3} = \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix} = \left(\begin{array}{cc|c} 26/32 & 5/32 & 1/32 \\ 16/32 & 10/32 & 6/32 \\ \hline 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|cc} Q_{2\times 2} & \underline{p}_{2\times 1} \\ \hline \underline{0}'_{1\times 2} & 1_{1\times 1} \end{array}\right)$$

where the essential transition probability sub-matrix $Q_{2\times 2}:(NA\to NA)$ is an $r\times r=2\times 2$ matrix, $\underline{p}_{2\times 1}:(NA\to A)$ is an $(r+s-1)\times 1=2\times 1$ column vector, $\underline{0}'_{1\times 2}:(A\to NA)$ is a $1\times(r+s-1)=1\times 2$ row vector and $1_{1\times 1}:(A\to A)$ represents the scalar value one.

The one-step transition probabilities are calculated by substituting SN_t in expression (2.46) by $2T-n$ and substituting in values for h, k, S_t^+ and S_{t-1}^+. The calculation of the one-step transition probabilities are given in Table 2.20 for illustration.

The probabilities in the last column of the TPM can also be calculated using the fact that $\sum_{j\in\Omega} p_{ij} = 1 \ \ \forall i$ (see equation (2.18)). Therefore,

* Note: Since only the state space needs to be described, S_{t-1}^+ can be any value from Ω^+ and we therefore take, without loss of generality, $S_{t-1}^+ = 0$. Any other possible value for S_{t-1}^+ would lead to the same Ω^+.

$p_{02} = 1-(p_{00}+p_{01}) = 1-(26/32+5/32) = 1/32$;

$p_{12} = 1-(p_{10}+p_{11}) = 1-(16/32+10/32) = 6/32$; and

$p_{22} = 1-(p_{20}+p_{21}) = 1-(0+0) = 1$.

Since it is easier to calculate the probabilities in the last column of the TPM using the latter approach, it will be used throughout the text from this point forward.

Table 2.20. The calculation of the transition probabilities when $h=4$, $k=1$ and $n=5$.

p_{00} $= P(S_t = 0 \mid S_{t-1} = 0)$ $= P(\min\{4, \max\{0, 0+SN_t - 1\}\} = 0)$ $= P(\max\{0, 0+SN_t - 1\} = 0)$ $= P(SN_t - 1 \le 0)$ $= P(SN_t \le 1)$ $= P(2T-5 \le 1)$ $= P(T \le 3)$ $= 26/32$	p_{01} $= P(S_t = 2 \mid S_{t-1} = 0)$ $= P(\min\{4, \max\{0, 0+SN_t - 1\}\} = 2)$ $= P(\max\{0, SN_t - 1\} = 2)$ $= P(SN_t - 1 = 2)$ $= P(SN_t = 3)$ $= P(2T-5 = 3)$ $= P(T = 4)$ $= 5/32$	p_{02} $= P(S_t = 4 \mid S_{t-1} = 0)$ $= P(\min\{4, \max\{0, 0+SN_t - 1\}\} = 4)$ $= P(\max\{0, SN_t - 1\} \ge 4)$ $= P(SN_t - 1 \ge 4)$ $= P(SN_t \ge 5)$ $= P(2T-5 \ge 5)$ $= 1-P(T \le 4)$ $= 1/32$
p_{10} $= P(S_t = 0 \mid S_{t-1} = 2)$ $= P(\min\{4, \max\{0, 2+SN_t - 1\}\} = 0)$ $= P(\max\{0, SN_t + 1\} = 0)$ $= P(SN_t \le -1)$ $= P(2T-5 \le -1)$ $= P(T \le 2)$ $= 16/32$	p_{11} $= P(S_t = 2 \mid S_{t-1} = 2)$ $= P(\min\{4, \max\{0, 2+SN_t - 1\}\} = 2)$ $= P(\max\{0, SN_t + 1\} = 2)$ $= P(SN_t = 1)$ $= P(2T-5 = 1)$ $= P(T = 3)$ $= 10/32$	p_{12} $= P(S_t = 4 \mid S_{t-1} = 2)$ $= P(\min\{4, \max\{0, 2+SN_t - 1\}\} = 4)$ $= P(\max\{0, SN_t + 1\} \ge 4)$ $= P(SN_t \ge 3)$ $= P(2T-5 \ge 3)$ $= 1-P(T \le 3)$ $= 6/32$
p_{20} $= P(S_t = 0 \mid S_{t-1} = 4)$ $= 0$ *	p_{21} $= P(S_t = 2 \mid S_{t-1} = 4)$ $= 0$	p_{22} $= P(S_t = 4 \mid S_{t-1} = 4)$ $= 1$ †

Using the TPM the *ARL* can be calculated using $ARL = \underline{\xi}(I-Q)^{-1}\underline{1}$. A well-known concern is that important information about the performance of a control chart can be missed when only examining the *ARL* (this is especially true when the process distribution is skewed). Various authors, see for example, Radson and Boyd (2005) and Chakraborti (2007), have suggested that one should examine a number of percentiles, including the median, to get the complete information about the performance of a control chart. Therefore, we now also consider percentiles. The $100\rho^{th}$ percentile is defined as the smallest integer l such that the

* The probability equals zero, because it is impossible to go from an absorbent state to a non-absorbent state.
† The probability equals one, since the probability of going from an absorbent state to an absorbent state is equal to one (once an absorbent state is entered, it is never left).

cdf is at least $(100\times\rho)\%$. Thus, the $100\rho^{th}$ percentile l is found from $P(N \le l) \ge \rho$. The median (50^{th} percentile) will be considered, since it is a more representative performance measure than the *ARL* (see the discussion in Section 2.1.5). The first and third quartiles (25^{th} and 75^{th} percentiles) will also be considered, since it contains the middle half of the distribution. The 'tails' of the distribution should also be examined and therefore the 5^{th} and 95^{th} percentiles are calculated. The calculation of these percentiles is shown in Table 2.21 for illustration purposes. The first column of Table 2.21 contains the values that the run length variable (N) can take on.

Table 2.21. Calculation of the percentiles when $h = 4$, $k = 1$ and $n = 5$[*].

N	$P(N \le l)$	The 5th, 25th, 50th, 75th and 95th percentiles
1	0.0313	
2	0.0859	$\rho_{0.05}$ = 2 (smallest integer such that the cdf is at least 0.05)
3	0.1420	
4	0.1954	
5	0.2456	
6	0.2928	$\rho_{0.25}$ = 6 (smallest integer such that the cdf is at least 0.25)
7	0.3370	
8	0.3784	
9	0.4173	
10	0.4537	
11	0.4878	
12	0.5198	$\rho_{0.5}$ = 12 (smallest integer such that the cdf is at least 0.5)
13	0.5499	
14	0.5780	
15	0.6044	
16	0.6291	
17	0.6523	
18	0.6740	
19	0.6944	
20	0.7135	
21	0.7314	
22	0.7482	
23	0.7639	$\rho_{0.75}$ = 23 (smallest integer such that the cdf is at least 0.75)
24	0.7787	
25	0.7925	
26	0.8055	
27	0.8176	
28	0.8290	
29	0.8397	
30	0.8497	
⋮	⋮	
48	0.9530	$\rho_{0.95}$ = 48 (smallest integer such that the cdf is at least 0.95)
49[†]	0.9559	

[*] See SAS Program 2 in Appendix B for the calculation of the values in Table 2.21.
[†] The value of the run length variable is only shown for some values up to *N*=49 for illustration purposes.

The formulas of the moments and some characteristics of the run length distribution have been studied by Fu, Spiring and Xie (2002) and Fu and Lou (2003) – see equations (2.41) to (2.45). By substituting $\underset{\sim}{\xi}_{1\times 2} = (1 \quad 0)$, $Q_{2\times 2} = \frac{1}{32}\begin{pmatrix} 26 & 5 \\ 16 & 10 \end{pmatrix}$ and $\underset{\sim}{1}_{2\times 1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ into these equations, we obtain the following:

$$ARL = E(N) = \underset{\sim}{\xi}(I-Q)^{-1}\underset{\sim}{1} = 16.62$$

$$E(N^2) = \underset{\sim}{\xi}(I+Q)(I-Q)^{-2}\underset{\sim}{1} = 516.59$$

$$SDRL = \sqrt{Var(N)} = \sqrt{E(N^2) - (E(N))^2} = 15.51$$

5^{th} percentile $= \rho_{0.05} = 2$

25^{th} percentile $= \rho_{0.25} = 6$

Median $= 50^{th}$ percentile $= \rho_{0.5} = 12$

75^{th} percentile $= \rho_{0.75} = 23$

95^{th} percentile $= \rho_{0.95} = 48$

Other values of *h*, *k* and *n* were also considered and the results are given in Table 2.22.

Table 2.22. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile* values for the upper one-sided CUSUM sign chart when $n = 5$†.

k	***h***	
	2	**3 or 4**
1	5.33 4.81 (1, 2, 4, 7, 15)	16.62 15.51 (2, 6, 12, 23, 48)
3	32.00 31.50 (2, 10, 22, 44, 95)	‡

* The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.
† See SAS Program 2 in Appendix B for the calculation of the values in Table 2.22.
‡ Since the decision interval is taken to satisfy $h \le n - k$ there are open cells in Table 2.22.

Note that the summary measures for odd values of h will be equal to the summary measures of the subsequent even integer. More on this later (refer to example (2.8)). Values of k and h are restricted to be integers so that the Markov chain approach could be employed to obtain expressions for the exact run length distribution and its parameters. In order to allow for the possibility of stopping after one sample, i.e. issuing a signal, the values of h is taken to satisfy $h \leq n-k$.

The five percentiles (given in Table 2.22) are displayed in boxplot-like graphs for various h and k values in Figure 2.5. It should be noted that these boxplot-like graphs differ from standard box plots. In the latter case the whiskers are drawn from the ends of the box to the smallest and largest values inside specified limits, whereas, in the case of the boxplot-like graphs, the whiskers are drawn from the ends of the box to the 5^{th} and 95^{th} percentiles, respectively. In this thesis "boxplot" will refer to a boxplot-like graph from this point forward.

Figure 2.5 clearly shows the effects of h and k on the run length distribution and it portrays the run length distribution when the process is in-control. We would prefer a "boxplot" with a high valued (large) in-control average run length and a small spread. Applying this criterion, we see that the "boxplot" corresponding to the $(h,k)=(2,3)$ combination has the largest in-control average run length, which is favorable, but it also has the largest spread which is unattractive. The "boxplot" furthest to the right is exactly opposite from the "boxplot" furthest to the left. The latter has the smallest spread, which is favorable, but it also has the smallest in-control average run length, which is unattractive. In conclusion, no "boxplot" is optimal relative to the others.

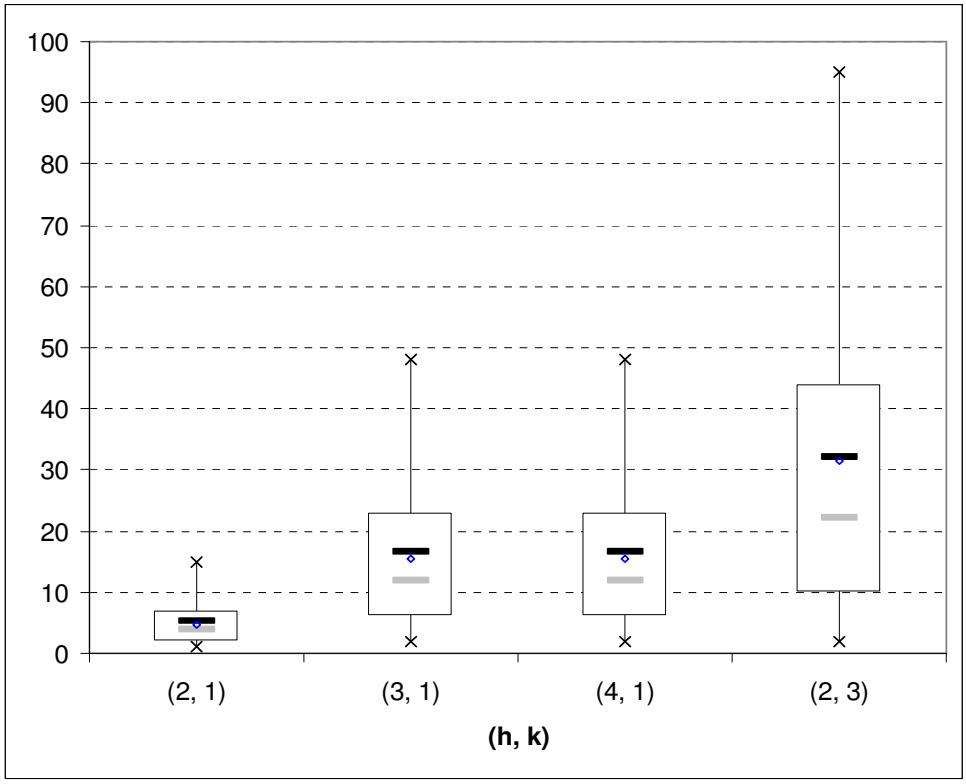

Figure 2.5. Boxplot-like graphs for the in-control run length distribution of various upper one-sided CUSUM sign charts when $n=5$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

Example 2.7

An upper one-sided CUSUM sign chart where the sample size is even (*n*=6)

The statistical properties of an upper one-sided CUSUM sign chart with a decision interval of 4 ($h=4$), a reference value of 2 ($k=2$) and a sample size of 6 ($n=6$) is examined. For n even, the reference value is taken to be even, because this leads to the sum $\sum(SN_i-k)$ being equal to even values which reduces the size of the state space for the Markov chain. For $h=4$ we have that the state space is $\Omega^+=\{\varsigma_0,\varsigma_1,\varsigma_2\}=\{0,2,4\}$ with $0=\varsigma_0<\varsigma_1<\varsigma_2=h$. The state space is calculated using equation (2.46) and the calculations are shown in Table 2.23.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Table 2.23. Calculation of the state space when $h=4$, $k=2$ and $n=6$.

SN_t	$S_{t-1}^+ + SN_t - k$	$\max\{0, S_{t-1}^+ + SN_t - k\}$	$S_t^+ = \min\{h, \max\{0, S_{t-1}^+ + SN_t - k\}\}$
-6	-8*	0	0
-4	-6	0	0
-2	-4	0	0
0	-2	0	0
2	0	0	0
4	2	2	2
6	4	4	4

Table 2.24. Classifications and descriptions of the states.

State number	Description of the state	Absorbent (A)/ Non-absorbent (NA)
0	$S_t^+ = 0$	NA
1	$S_t^+ = 2$	NA
2	$S_t^+ = 4$	A

From Table 2.24 we see that there are two non-absorbent states, i.e. $r=2$, and one absorbent state, i.e. $s=1$. Therefore, the corresponding TPM will be a $(r+s)\times(r+s)=3\times 3$ matrix. It can be shown (see Table 2.25) that the TPM is given by

$$TPM_{3\times 3} = \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix} = \left(\begin{array}{cc|c} 57/64 & 6/64 & 1/64 \\ 42/64 & 15/64 & 7/64 \\ \hline 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|c|c} Q_{2\times 2} & & \underline{p}_{2\times 1} \\ - & - & - \\ \underline{0}'_{1\times 2} & & 1_{1\times 1} \end{array}\right)$$

where the essential transition probability sub-matrix $Q_{2\times 2}:(NA\to NA)$ is an $r\times r=2\times 2$ matrix, $\underline{p}_{2\times 1}:(NA\to A)$ is an $(r+s-1)\times 1=2\times 1$ column vector, $\underline{0}'_{1\times 2}:(A\to NA)$ is a $1\times(r+s-1)=1\times 2$ row vector and $1_{1\times 1}:(A\to A)$ represents the scalar value one. The one-step transition probabilities are calculated by substituting SN_t in expression (2.46) by $2T-n$ and substituting in values for h, k, S_t^+ and S_{t-1}^+. The calculation of the one-step transition probabilities are given in Table 2.25 for illustration.

* Note: Since only the state space needs to be described, S_{t-1}^+ can be any value from Ω^+ and we therefore take, without loss of generality, $S_{t-1}^+ = 0$. Any other possible value for S_{t-1}^+ would lead to the same Ω^+.

Table 2.25. The calculation of the transition probabilities when $h = 4$, $k = 2$ and $n = 6$.

p_{00} $= P(S_t = 0 \mid S_{t-1} = 0)$ $= P(\min\{4, \max\{0, 0 + SN_t - 2\}\} = 0)$ $= P(\max\{0, 0 + SN_t - 2\} = 0)$ $= P(SN_t - 2 \le 0)$ $= P(SN_t \le 2)$ $= P(T \le 4)$ $= {}^{57}/_{64}$	p_{01} $= P(S_t = 2 \mid S_{t-1} = 0)$ $= P(\min\{4, \max\{0, 0 + SN_t - 2\}\} = 2)$ $= P(\max\{0, SN_t - 2\} = 2)$ $= P(SN_t - 2 = 2)$ $= P(SN_t = 4)$ $= P(T = 5)$ $= {}^{6}/_{64}$	p_{02} $= 1 - (p_{00} + p_{01})$ $= 1 - ({}^{57}/_{64} + {}^{6}/_{64})$ $= {}^{1}/_{64}$
p_{10} $= P(S_t = 0 \mid S_{t-1} = 2)$ $= P(\min\{4, \max\{0, 2 + SN_t - 2\}\} = 0)$ $= P(\max\{0, SN_t\} = 0)$ $= P(SN_t \le 0)$ $= P(T \le 3)$ $= {}^{42}/_{64}$	p_{11} $= P(S_t = 2 \mid S_{t-1} = 2)$ $= P(\min\{4, \max\{0, 2 + SN_t - 2\}\} = 2)$ $= P(\max\{0, SN_t\} = 2)$ $= P(SN_t = 2)$ $= P(T = 4)$ $= {}^{15}/_{64}$	p_{12} $= 1 - (p_{10} + p_{11})$ $= 1 - ({}^{42}/_{64} + {}^{15}/_{64})$ $= {}^{7}/_{64}$
p_{20} $= P(S_t = 0 \mid S_{t-1} = 4)$ $= 0$ *	p_{21} $= P(S_t = 2 \mid S_{t-1} = 4)$ $= 0$	p_{22} $= 1 - (p_{20} + p_{21})$ $= 1$

The formulas of the moments and some characteristics of the run length distribution have been studied by Fu, Spiring and Xie (2002) and Fu and Lou (2003) – see equations (2.41) to (2.45). By substituting $\underline{\xi}_{1\times 2} = (1 \quad 0)$, $Q_{2\times 2} = \frac{1}{64}\begin{pmatrix} 57 & 6 \\ 42 & 15 \end{pmatrix}$ and $\underline{1}_{2\times 1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ into these equations, we obtain the following:

$$ARL = E(N) = \underline{\xi}(I - Q)^{-1}\underline{1} = 38.68$$

$$E(N^2) = \underline{\xi}(I + Q)(I - Q)^{-2}\underline{1} = 2918.19$$

$$SDRL = \sqrt{Var(N)} = \sqrt{E(N^2) - (E(N))^2} = 37.71$$

5^{th} percentile $= \rho_{0.05} = 3$

25^{th} percentile $= \rho_{0.25} = 12$

Median $= 50^{th}$ percentile $= \rho_{0.5} = 27$

75^{th} percentile $= \rho_{0.75} = 53$

95^{th} percentile $= \rho_{0.95} = 114$

* The probability equals zero, because it is impossible to go from an absorbent state to a non-absorbent state.

Other values of *h*, *k* and *n* were also considered and the results are given in Table 2.26.

Table 2.26. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile* values for the upper one-sided CUSUM sign chart when $n = 6$†.

k	***h***		
	2	**3 or 4**	**5 or 6**
0	2.91 2.36 (1, 1, 2, 4, 8)	5.92 4.96 (1, 2, 4, 8, 16)	10.66 8.80 (2, 4, 8, 14, 28)
2	9.14 8.63 (1, 3, 6, 12, 26)	38.68 37.71 (3, 12, 27, 53, 114)	‡
4	64.00 63.50 (4, 19, 45, 89, 191)		

The five percentiles (given in Table 2.26) are displayed in boxplot-like graphs for various h and k values in Figure 2.6. Recall that we would prefer a "boxplot" with a high valued (large) in-control average run length and a small spread. Applying this criterion, we see that the "boxplot" corresponding to the $(h,k) = (2,4)$ combination has the largest in-control average run length, which is favorable, but it also has the largest spread which is unattractive. The "boxplot" furthest to the right is exactly opposite from the "boxplot" furthest to the left. The latter has the smallest spread, which is favorable, but it also has the smallest in-control average run length, which is unattractive. In conclusion, no "boxplot" is optimal relative to the others.

* The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.
† See SAS Program 2 in Appendix B for the calculation of the values in Table 2.26.
‡ Since the decision interval is taken to satisfy $h \le n - k$ there are open cells in Table 2.26.

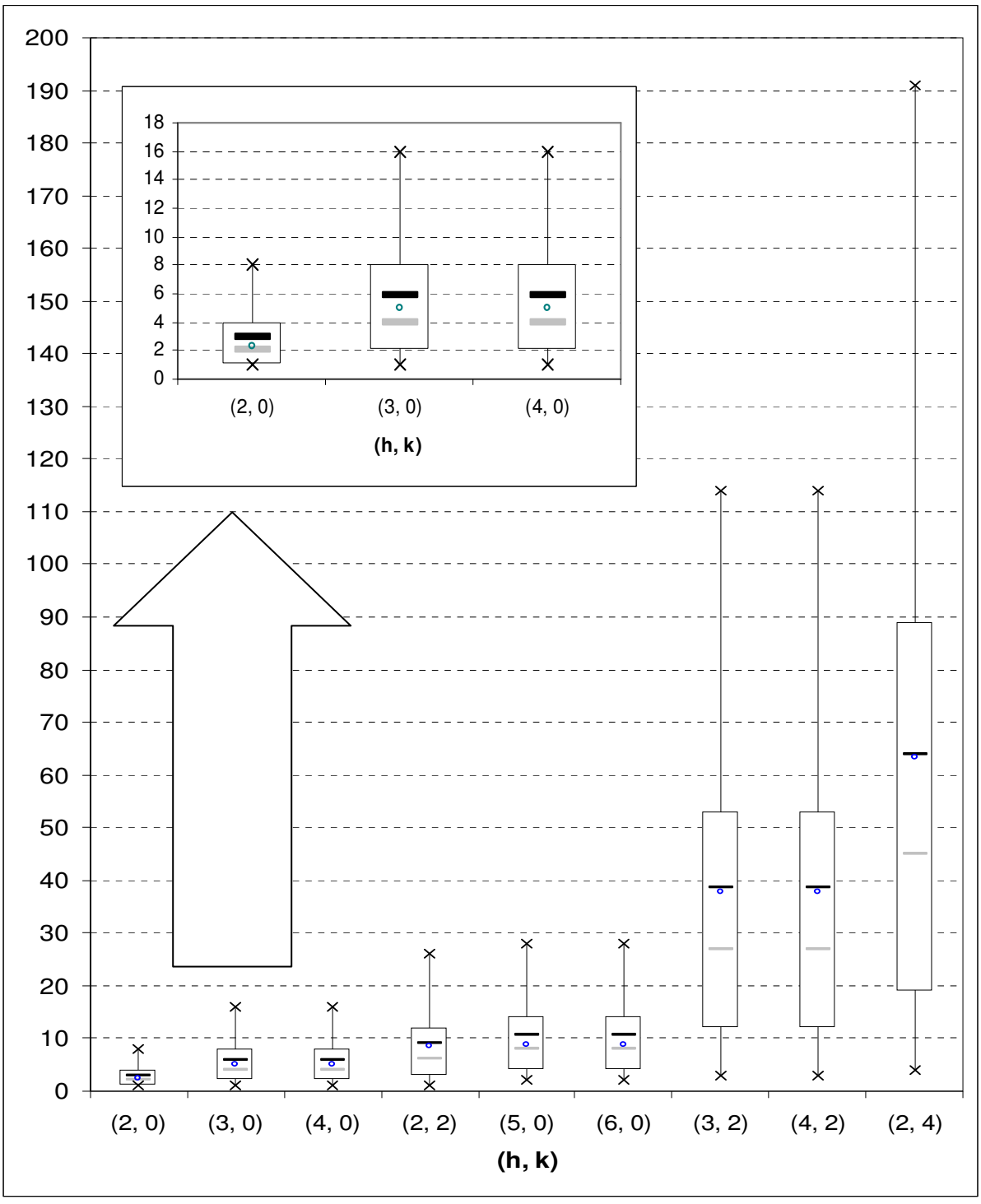

Figure 2.6. Boxplot-like graphs for the in-control run length distribution of various upper one-sided CUSUM sign charts when $n = 6$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "○" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

On the performance side, note that the largest in-control average run length that the upper one-sided CUSUM sign chart can obtain is 2^n. Therefore, for a sample size of 6 the largest ARL_0^+ equals $2^6 = 64$ (this is obtained when $h = 2$ and $k = 4$). For this case we find the $TPM = \begin{pmatrix} 63/64 & 1/64 \\ 0 & 1 \end{pmatrix}$ and as a result the in-control average run length equals $ARL_0^+ = \underline{\xi}(I - Q)^{-1}\underline{1} = 1 \times (1 - 63/64)^{-1} \times 1 = 64$. Since the largest ARL_0^+ is only 64 for $n = 6$, many false alarms will be expected by this chart leading to a possible loss of time and resources. Larger sample sizes should therefore preferably be taken when implementing the upper one-sided CUSUM sign chart.

.....

Example 2.8

An upper one-sided CUSUM sign chart with a decision interval of 4 (*h*=4), a reference value of 1 (*k*=1) and a sample size of 5 (*n*=5)

In the previous two examples it can be seen that summary measures for odd values of h will be equal to the summary measures of the subsequent even integer. This will be illustrated by the use of an example.

For the upper one-sided CUSUM sign chart with a decision interval of 4 ($h = 4$), a reference value of 1 ($k = 1$) and a sample size of 5 ($n = 5$) the TPM was given by $TPM = \begin{pmatrix} 26/32 & 5/32 & 1/32 \\ 16/32 & 10/32 & 6/32 \\ 0 & 0 & 1 \end{pmatrix}$ (see example (2.6). By keeping the reference value and the sample size fixed and changing h to an odd integer ($h = 3$) we obtain the same TPM and therefore we obtain the same summary measures. Stated differently, the summary measures of h odd ($h = 3$) will be equal to the summary measures of the subsequent even integer ($h = 4$).

.....

We've considered sample sizes of $n = 5$ and 6 and established that larger sample sizes should preferably be taken when implementing the upper one-sided CUSUM sign chart. Therefore, a larger sample size ($n = 10$) is considered and the results are given in Table 2.27.

Table 2.27. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile* values for the upper one-sided CUSUM sign chart when $n = 10$†.

k	*h*		
	3 or 4	**5 or 6**	**7 or 8**
2	14.34 13.58 (1, 5, 10, 20, 41)	36.81 35.48 (3, 12, 26, 51, 108)	91.59 89.45 (7, 28, 64, 126, 270)
4	77.97 77.29 (5, 23, 54, 108, 232)	464.86 463.68 (25, 135, 323, 644, 1390)	‡
6	929.97 929.37 (48, 268, 645, 1289, 2785)		

Table 2.27 gives values of ARL_0^+ for various values of h and k when the sample size is equal to 10. Amin, Bakir and Reynolds (1995) provided a similar Table (see Table 5 on page 1613) containing the in-control run length summary values for the upper one-sided CUSUM sign chart (ARL_0^+) for a range of k and h values when $n = 10$.

The five percentiles (given in Table 2.27) are displayed in boxplot-like graphs for various h and k values in Figure 2.7. Recall that we would prefer a "boxplot" with a high valued (large) in-control average run length and a small spread. Applying this criterion, we see that the "boxplot" corresponding to the $(h,k) = (3,6)$ or $(h,k) = (4,6)$ combination has the largest in-control average run length, which is favorable, but it also has the largest spread which is unattractive. The "boxplot" furthest to the right is exactly opposite from the "boxplot" furthest to the left. The latter has the smallest spread, which is favorable, but it also has the smallest in-control average run length, which is unattractive. In conclusion, no "boxplot" is optimal relative to the others.

* The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.
† See SAS Program 2 in Appendix B for the calculation of the values in Table 2.27.
‡ Since the decision interval is taken to satisfy $h \le n - k$ there are open cells in Table 2.27.

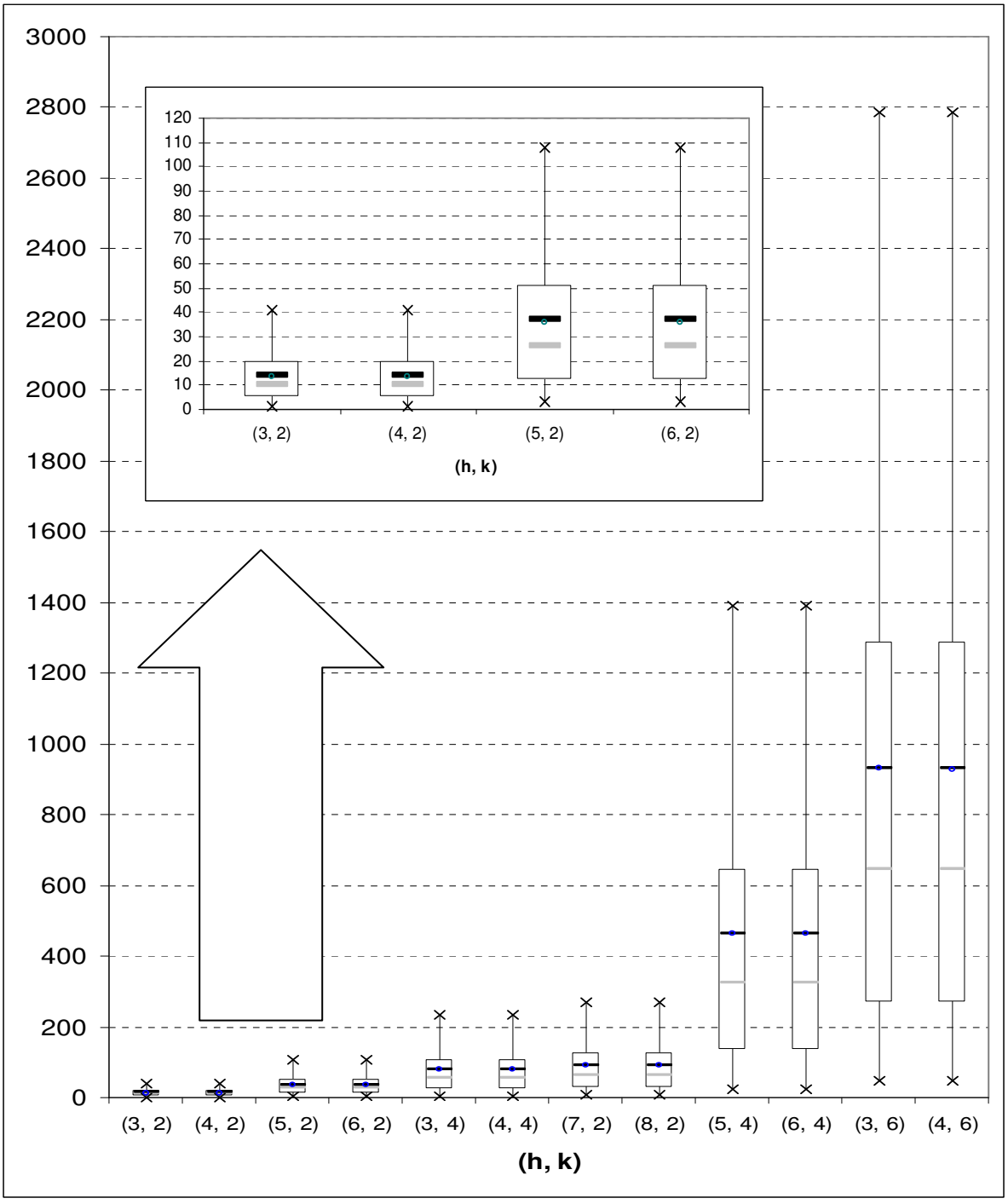

Figure 2.7. Boxplot-like graphs for the in-control run length distribution of various upper one-sided CUSUM sign charts when $n = 10$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "○" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Example 2.9

An upper one-sided CUSUM sign chart for the Montgomery (2001) piston ring data

We conclude this sub-section by illustrating the upper one-sided CUSUM sign chart using a set of data from Montgomery (2001; Table 5.2) on the inside diameters of piston rings manufactured by a forging process. The dataset contains 15 samples (each of size 5). We assume that the underlying distribution is symmetric with a known target value of $\theta_0 = 74\ mm$.

Let $k = 3$. Once k is selected, the constant h should be chosen to give the desired in-control average run length performance. By choosing $h = 2$ we obtain an in-control average run length of 32 which is the highest in-control average run length attainable when $n = 5$ (see Table 2.22).

The plotting statistics for the Shewhart sign chart (SN_i for $i = 1,2,...,15$) are given in the second row of Table 2.28. The upper one-sided CUSUM plotting statistics (S_i^+ for $i = 1,2,...,15$) are given in the third row of Table 2.28.

Table 2.28. SN_i and S_i^+ values for the piston ring data in Montgomery (2001)*.

Sample No:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
SN_i	2	1	-4	3	0	3	3	-1	3	4	1	5	5	5	4
S_i^+	0	0	0	0	0	0	0	0	0	1	0	2	4	6	7

To illustrate the calculations, consider sample number 1. The equation for the plotting statistic S_i^+ is $S_1^+ = \max[0, S_0^+ + SN_1 - k] = \max[0,0+2-3] = \max[0,-1] = 0$ where a signaling event occurs for the first i such that $S_i^+ \geq h$, that is, $S_i^+ \geq 2$. The graphical display of the upper one-sided CUSUM sign chart is shown in Figure 2.8.

* See SAS Program 3 in Appendix B for the calculation of the values in Table 2.28.

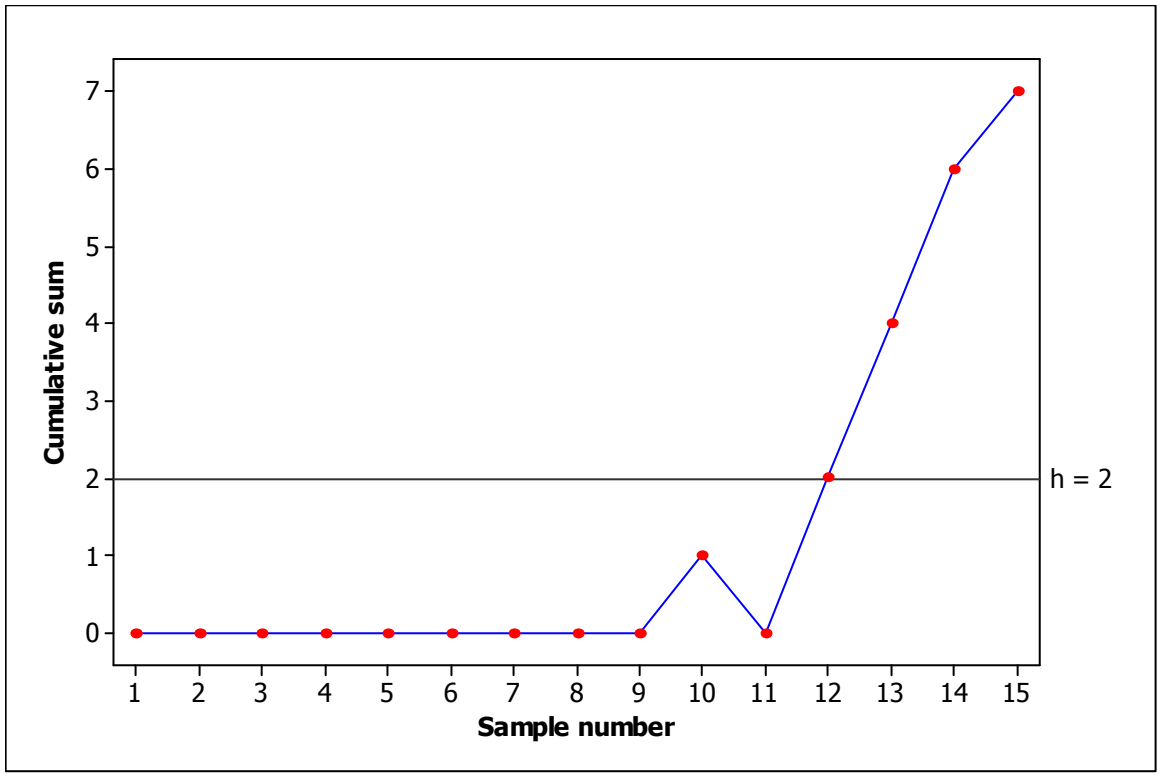

Figure 2.8. The upper one-sided CUSUM sign chart for the Montgomery (2001) piston ring data.

The upper one-sided CUSUM sign chart signals at sample 12, indicating a most likely upward shift from the known target value θ_0. The action taken following an out-of-control signal on a CUSUM chart is identical to that with any control chart. A search for assignable causes should be done, corrective action should be taken (if required) and, following this, the CUSUM is reset to zero. Different control charts are compared by designing the control charts to have the same ARL_0 and then evaluating the ARL_δ. The control chart with the lower ARL_δ is the preferred chart. These procedures can not be meaningfully illustrated using the data from Montgomery (2001) because the sample size $n = 5$ used here is too small. It may be noted that the highest ARL_0^+ is 32 for $n = 5$. Thus, achievable values of ARL_0^+ are too small for practical use, unless *n* is 'large'.

2.3.2.2. Lower one-sided control charts

Analogous to the previous section, a Markov chain representation of the one-sided CUSUM procedure based on integer-valued cumulative sums is presented in this section. The number of states included in the Markov chain is minimized in order to make the methods as efficient as possible. The time that the procedure signals is the first time such that the finite-state Markov chain S_t^- enters the state ς_0 where the state space is given by $\Omega^- = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$ with $-h = \varsigma_0 < ... < \varsigma_{r+s-1} = 0$, $S_0^- = 0$ and

$$S_t^- = \max\left\{-h, \min\left\{0, S_{t-1}^- + SN_t + k\right\}\right\}. \tag{2.47}$$

Clearly, there is only one absorbent state, since the chart signals when S_t^- falls on or below $-h$, i.e. $s = 1$.

The distribution of SN_t can easily be obtained from the binomial distribution (recall that $SN_i = 2T_i - n \quad \forall i$, where T_i is binomially distributed with parameters n and $p = P(X_{ij} \geq \theta_0))$. The binomial probabilities are given in Table G of Gibbons and Chakraborti (2003) and can easily be calculated using some type of statistical software package, for example, Excel or SAS.

Example 2.10

A lower one-sided CUSUM sign chart where the sample size is odd ($n = 5$)

The statistical properties of a lower one-sided CUSUM sign chart with a decision interval of 4 ($h = 4$), a reference value of 1 ($k = 1$) and a sample size of 5 ($n = 5$) is examined. For n odd, the reference value is taken to be odd, because this leads to the sum $\sum(SN_i - k)$ being equal to even values which reduces the size of the state space for the Markov chain. For $h = 4$ we have $\Omega^- = \{\varsigma_0, \varsigma_1, \varsigma_2\} = \{-4, -2, 0\}$ with $-h = \varsigma_0 < \varsigma_1 < \varsigma_2 = 0$. The state space is calculated using equation (2.47) and the calculations are shown in Table 2.29.

Table 2.29. Calculation of the state space when $h=4$, $k=1$ and $n=5$.

SN_t	$S_{t-1}^- + SN_t + k$	$\min\{0, S_{t-1}^- + SN_t + k\}$	$S_t^- = \max\{-h, \min\{0, S_{t-1}^- + SN_t + k\}\}$
-5	-4*	-4	-4
-3	-2	-2	-2
-1	0	0	0
1	2	0	0
3	4	0	0
5	6	0	0

Table 2.30. Classifications and descriptions of the states.

State number	Description of the state	Absorbent (A)/ Non-absorbent (NA)
0	$S_t^- = 0$	NA
1	$S_t^- = -2$	NA
2	$S_t^- = -4$	A

From Table 2.30 we see that there are two non-absorbent states, i.e. $r=2$, and one absorbent state, i.e. $s=1$. Therefore, the corresponding TPM will be a $(r+s)\times(r+s)=3\times 3$ matrix. It can be shown (see Table 2.31) that the TPM is given by

$$TPM_{3\times 3} = \begin{pmatrix} p_{00} & p_{0(-2)} & p_{0(-4)} \\ p_{(-2)0} & p_{(-2)(-2)} & p_{(-2)(-4)} \\ p_{(-4)0} & p_{(-4)(-2)} & p_{(-4)(-4)} \end{pmatrix} = \left(\begin{array}{cc|c} 26/32 & 5/32 & 1/32 \\ 16/32 & 10/32 & 6/32 \\ \hline 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} Q_{2\times 2} & \underline{p}_{2\times 1} \\ \hline \underline{0}'_{1\times 2} & 1_{1\times 1} \end{array}\right)$$

where the essential transition probability sub-matrix $Q_{2\times 2}:(NA\to NA)$ is an $r\times r=2\times 2$ matrix, $\underline{p}_{2\times 1}:(NA\to A)$ is an $(r+s-1)\times 1=2\times 1$ column vector, $\underline{0}'_{1\times 2}:(A\to NA)$ is a $1\times(r+s-1)=1\times 2$ row vector and $1_{1\times 1}:(A\to A)$ represents the scalar value one.

The one-step transition probabilities are calculated by substituting SN_t in expression (2.47) by $2T-n$ and substituting in values for h, k, S_t^- and S_{t-1}^-. The calculation of the one-step transition probabilities are given for illustration in Table 2.31.

* Note: Since only the state space needs to be described, S_{t-1}^- can be any value from Ω^- and we therefore take, without loss of generality, $S_{t-1}^- = 0$. Any other possible value for S_{t-1}^- would lead to the same Ω^-.

Table 2.31. The calculation of the transition probabilities when $h = 4$, $k = 1$ and $n = 5$.

p_{00} $= P(S_t = 0 \mid S_{t-1} = 0)$ $= P(\max\{-4, \min\{0,0 + SN_t + 1\}\} = 0)$ $= P(\min\{0,0 + SN_t + 1\} = 0)$ $= P(SN_t + 1 \geq 0)$ $= P(SN_t \geq -1)$ $= P(2T - 5 \geq -1)$ $= P(T \geq 2)$ $= 1 - P(T \leq 1)$ $= {}^{26}/_{32}$	$p_{0(-2)}$ $= P(S_t = -2 \mid S_{t-1} = 0)$ $= P(\max\{-4, \min\{0,0 + SN_t + 1\}\} = -2)$ $= P(\min\{0, SN_t + 1\} = -2)$ $= P(SN_t + 1 = -2)$ $= P(SN_t = -3)$ $= P(2T - 5 = -3)$ $= P(T = 1)$ $= {}^{5}/_{32}$	$p_{0(-4)}$ $= 1 - (p_{00} + p_{0(-2)})$ $= 1 - ({}^{26}/_{32} + {}^{5}/_{32})$ $= {}^{1}/_{32}$
$p_{(-2)0}$ $= P(S_t = 0 \mid S_{t-1} = -2)$ $= P(\max\{-4, \min\{0,-2 + SN_t + 1\}\} = 0)$ $= P(\min\{0,-2 + SN_t + 1\} = 0)$ $= P(SN_t - 1 \geq 0)$ $= P(SN_t \geq 1)$ $= P(2T - 5 \geq 1)$ $= P(T \geq 3)$ $= 1 - P(T \leq 2)$ $= {}^{16}/_{32}$	$p_{(-2)(-2)}$ $= P(S_t = -2 \mid S_{t-1} = -2)$ $= P(\max\{-4, \min\{0,-2 + SN_t + 1\}\} = -2)$ $= P(\min\{0,-2 + SN_t + 1\} = -2)$ $= P(SN_t - 1 = -2)$ $= P(SN_t = -1)$ $= P(2T - 5 = -1)$ $= P(T = 2)$ $= {}^{10}/_{32}$	$p_{(-2)(-4)}$ $= 1 - (p_{(-2)0} + p_{(-2)(-2)})$ $= 1 - ({}^{16}/_{32} + {}^{10}/_{32})$ $= {}^{6}/_{32}$
$p_{(-4)0}$ $= P(S_t = 0 \mid S_{t-1} = -4)$ $= 0$ *	$p_{(-4)(-2)}$ $= P(S_t = -2 \mid S_{t-1} = -4)$ $= 0$	$p_{(-4)(-4)}$ $= 1 - (p_{(-4)0} + p_{(-4)(-2)})$ $= 1 - (0 + 0)$ $= 1$

The formulas of the moments and some characteristics of the run length distribution have been studied by Fu, Spiring and Xie (2002) and Fu and Lou (2003) – see equations (2.41) to (2.45). By substituting $\underline{\xi}_{1\times 2} = (1 \quad 0)$, $Q_{2\times 2} = \begin{pmatrix} {}^{26}/_{32} & {}^{5}/_{32} \\ {}^{16}/_{32} & {}^{10}/_{32} \end{pmatrix}$ and $\underline{1}_{2\times 1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ into these equations, we obtain the following:

$$ARL = E(N) = \underline{\xi}(I - Q)^{-1}\underline{1} = 16.62$$

$$E(N^2) = \underline{\xi}(I + Q)(I - Q)^{-2}\underline{1} = 516.59$$

$$SDRL = \sqrt{Var(N)} = \sqrt{E(N^2) - (E(N))^2} = 15.51$$

5^{th} percentile $= \rho_5 = 2$

25^{th} percentile $= \rho_{25} = 6$

* The probability equals zero, because it is impossible to go from an absorbent state to a non-absorbent state.

$\text{Median} = 50^{th}\ \text{percentile} = \rho_{50} = 12$

$75^{th}\ \text{percentile} = \rho_{75} = 23$

$95^{th}\ \text{percentile} = \rho_{95} = 48$

The in-control average run length values, standard deviation of the run length values and percentiles for the lower one-sided CUSUM sign chart are exactly the same as for the upper one-sided CUSUM sign chart, since the one-step transition probabilities matrices are the same (compare the transition probabilities matrices of examples 2.6 and 2.10). Therefore, we obtain Result 2.10:

Result 2.10:

The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile values tabulated for the upper one-sided CUSUM sign chart will also hold for the lower one-sided CUSUM sign chart.

Example 2.11

A lower one-sided CUSUM sign chart for the Montgomery (2001) piston ring data

We conclude this sub-section by illustrating the lower one-sided CUSUM sign chart using a set of data from Montgomery (2001; Table 5.2) on the inside diameters of piston rings manufactured by a forging process. The dataset contains 15 samples (each of size 5). We assume that the underlying distribution is symmetric with a known target value of $\theta_0 = 74\,mm$.

From Table 2.22 it can be seen that the in-control average run length equals 32 when $h = 2$ and $k = 3$ (recall that this is the largest possible in-control average run length value that the chart can obtain, since $2^5 = 32$). The plotting statistics for the Shewhart sign chart (SN_i for $i = 1,2,...,15$) are given in the second row of Table 2.32. The lower one-sided CUSUM plotting statistics (S_i^- for $i = 1,2,...,15$) are given in the third row of Table 2.32.

Table 2.32. SN_i and S_i^- values for the piston ring data in Montgomery (2001)*.

Sample No:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SN_i	2	1	-4	3	0	3	3	-1	3	4	1	5	5	5	4
S_i^-	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0

To illustrate the calculations, consider sample number 1. The equation for the plotting statistic S_1^- is

$$S_1^{-*} = \max[0, S_0^{-*} - SN_1 - k] = \max[0,0-2-3] = \max[0,-5] = 0 \qquad (2.48)$$

or

$$S_1^- = \min[0, S_0^- + SN_1 + k] = \min[0,0+2+3] = \min[0,5] = 0\,. \qquad (2.49)$$

A signaling event occurs for the first i such that $S_i^{-*} \geq h$, that is, $S_i^{-*} \geq 2$ if expression (2.48) is used or $S_i^- \leq -h$, that is, $S_i^- \leq -2$ if expression (2.49) is used.

The graphical display of the lower one-sided CUSUM sign chart (using expression (2.49)) is shown in Figure 2.9 and does not signal.

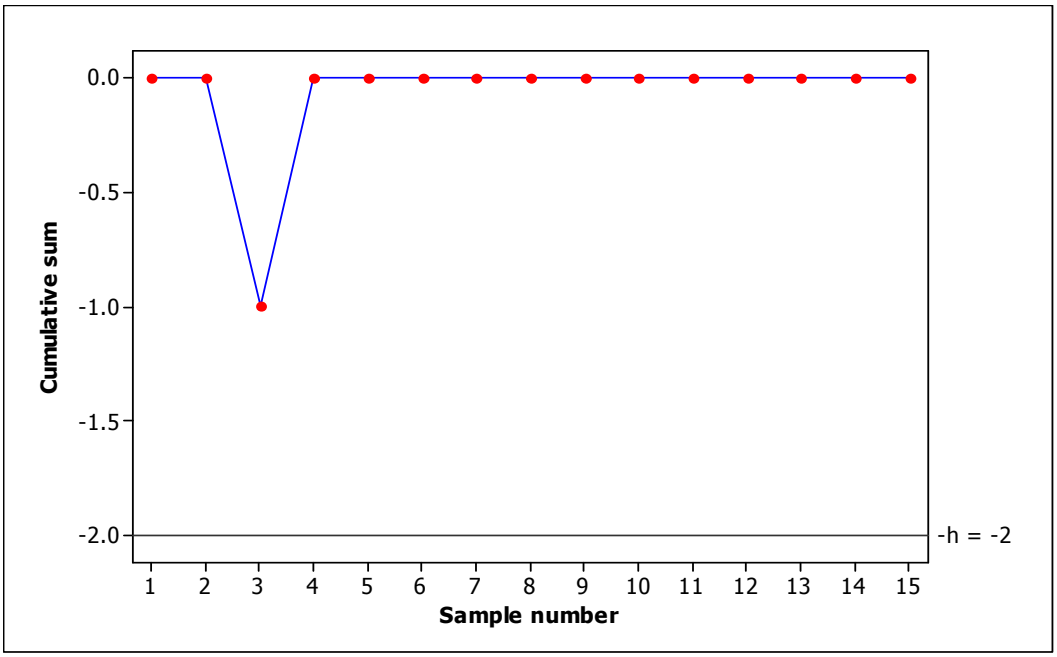

Figure 2.9. The lower one-sided CUSUM sign chart for the Montgomery (2001) piston ring data.

* See SAS Program 3 in Appendix B for the calculation of the values in Table 2.32.

2.3.3. Two-sided control charts

Various authors have studied the two-sided CUSUM scheme, for example, Kemp (1971) gives the average run length of the two-sided CUSUM scheme in terms of the average run lengths of the two corresponding one-sided schemes. Lucas and Crosier (1982) used a Markov chain representation of the two-sided CUSUM scheme to determine the run length distribution and the average run length. In this thesis, the approach taken by Brook and Evans (1972) for the one-sided CUSUM scheme is extended to the two-sided CUSUM scheme. A Markov chain representation of the two-sided CUSUM scheme based on integer-valued random variables will be presented. This is done since the nonparametric statistics that are the building blocks of the CUSUM scheme are discrete random variables. The number of states included in the Markov chain is minimized by taking the reference value k so that the state space of the Markov chain is a set of even numbers. This reduces the size of the TPM and thus eliminates unnecessary calculations in order to make the methods as efficient as possible.

Recall that for the upper one-sided CUSUM sign chart we use

$$S_t^+ = \min\left\{h, \max\{0, SN_t - k + S_{t-1}^+\}\right\} \text{ for } n = 1,2,... \tag{2.50}$$

For a lower one-sided CUSUM sign chart we use

$$S_t^- = \max\left\{-h, \min\{0, SN_t + k + S_{t-1}^-\}\right\} \text{ for } n = 1,2,... \tag{2.51}$$

For the two-sided scheme the two one-sided schemes are performed simultaneously. The corresponding two-sided CUSUM chart signals for the first t at which either one of the two inequalities is satisfied, that is, either $S_t^+ \geq h$ or $S_t^- \leq -h$. Starting values are typically chosen to equal zero, that is, $S_0^+ = S_0^- = 0$. The two-sided scheme signals at $N = \min_t\left\{t : S_t^+ \geq h \quad \text{or} \quad S_t^- \leq -h\right\}$ where h is a positive integer.

The two-sided CUSUM scheme can be represented by a Markov chain with states corresponding to the possible combinations of values of S_t^+ and S_t^-. The states corresponding to values for which a signal is given by the CUSUM scheme are called absorbent states. Clearly, there are two absorbent states since the chart signals when S_t^+ falls on or above h or when S_t^- falls on or below $-h$, i.e. $s = 2$. Recall that r denotes the number of non-absorbent

states and, consequently, the corresponding TPM is an $(r+s)\times(r+s)$, i.e. an $(r+2)\times(r+2)$ matrix.

The time that the procedure signals is the first time such that the finite-state Markov chain enters the state ς_0 or ς_{r+s-1} where the state space is given by $\Omega = \Omega^+ \cup \Omega^- = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$ with $-h = \varsigma_0 < ... < \varsigma_{r+s-1} = h$.

Example 2.12

A two-sided CUSUM sign chart where the sample size is odd ($n = 5$)

The statistical properties of a two-sided CUSUM sign chart with a decision interval of 4 ($h = 4$), a reference value of 1 ($k = 1$) and a sample size of 5 ($n = 5$) is examined. For n odd, the reference value k is taken to be odd, because this leads to the sum $\sum(SN_i - k)$ being equal to even values which reduces the size of the state space for the Markov chain. This reduces the size of the TPM and thus eliminates unnecessary calculations in order to make the methods as efficient as possible. Let Ω denote the state space for the two-sided chart. Ω is the union of the state space for the upper one-sided chart $\Omega^+ = \{0,2,4\}$ and the state space for the lower one-sided chart $\Omega^- = \{-4,-2,0\}$. Therefore, $\Omega = \Omega^+ \cup \Omega^- = \{-4,-2,0\} \cup \{0,2,4\} = \{-4,-2,0,2,4\} = \{\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4\}$ with $-h = \varsigma_0 < \varsigma_1 < \varsigma_2 < \varsigma_3 < \varsigma_4 = h$.

Table 2.33. Classifications and descriptions of the states.

State number	Values of the CUSUM statistic(s)	Absorbent (A) Non-absorbent (NA)
0	$S_t^- = 0$ and $S_t^+ = 0$	NA
1	$S_t^- = 2$ or $S_t^+ = 2$ *	NA
2	$S_t^- = -2$ or $S_t^+ = -2$ †	NA
3	$S_t^- = 4$ or $S_t^+ = 4$ ‡	A
4	$S_t^- = -4$ or $S_t^+ = -4$ §	A

From Table 2.33 we see that there are three non-absorbent states, i.e. $r = 3$, and two absorbent states, i.e. $s = 2$. Therefore, the corresponding TPM will be a $(r+s)\times(r+s) = 5\times 5$ matrix. The layout of the TPM is as follows. There are three transient states and two absorbent states. By convention we first list the non-absorbent states and then we list the absorbent states. In column one we compute the probability of moving from state i to state 0, for all i. Note that the process reaches state 0 when both the upper and the lower cumulative sums equal zero. In columns two and three, we compute the probabilities of moving from state i to the remaining non-absorbent states, for all i. Finally, in the remaining two columns we compute the probabilities of moving from state i to the absorbent states, for all i. Thus, the TPM can be conveniently partitioned into 4 sections given by

$$TPM_{5\times 5} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \left(\begin{array}{ccc|cc} 20/32 & 5/32 & 5/32 & 1/32 & 1/32 \\ 10/32 & 10/32 & 5/32 & 6/32 & 1/32 \\ 10/32 & 5/32 & 10/32 & 1/32 & 6/32 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} Q_{3\times 3} & C_{3\times 2} \\ \hline Z_{2\times 3} & I_{2\times 2} \end{array}\right)$$

* Moving from state 0 to state 1 can happen when either the upper cumulative sum or the lower cumulative sum equals 2. But the lower cumulative sum can not equal 2 since by definition the lower cumulative sum can only take on integer values smaller than or equal to zero. Therefore, we only use the probability that the upper cumulative sum equals 2 in the calculation of the probabilities in the TPM.

† Moving from state 0 to state 2 can happen when either the upper cumulative sum or the lower cumulative sum equals -2. But the upper cumulative sum can not equal -2 since by definition the upper cumulative sum can only take on integer values greater than or equal to zero. Therefore, we only use the probability that the lower cumulative sum equals -2 in the calculation of the probabilities in the TPM.

‡ A similar argument to the argument in the first footnote on this page holds. Therefore, we only use the probability that the upper cumulative sum equals 4 in the calculation of the probabilities in the TPM.

§ A similar argument to the argument in the second footnote on this page holds. Therefore, we only use the probability that the lower cumulative sum equals -4 in the calculation of the probabilities in the TPM.

where the essential transition probability sub-matrix $Q_{3\times3}:(NA \to NA)$ is an $r \times r = 3\times3$ matrix, $C_{3\times2}:(NA \to A)$ is an $r \times s = 3\times2$ matrix, $Z_{2\times3}:(A \to NA)$ is an $s \times r = 2\times3$ matrix and $I_{2\times2}:(A \to A)$ is a $s \times s = 2\times2$ matrix.

The calculation of the elements of the TPM is illustrated next. Note that this essentially involves the calculation of the matrices Q and C. First consider the transient states. Note that the process moves to state 0, i.e., $j = 0$, when both the upper *and* the lower cumulative sums equal 0. Thus the required probability of moving to 0, from any other transient state, is the probability of an *intersection* of two sets involving values of the upper and the lower CUSUM statistics, respectively. On the other hand, the probability of moving to any state $j \neq 0$, from any other state, is the probability of a *union* of two sets involving values of the upper and the lower CUSUM statistics, respectively. However, one of these two sets is empty so that the required probability is the probability of only the non-empty set.

The calculation of the entry in the first row and the first column of the matrix Q, p_{00}, will be discussed in detail. This is the probability of moving from state 0 to state 0 in one step. As we just described, this can happen only when the upper *and* the lower cumulative sums both equal 0 at time t. For the upper one-sided CUSUM p_{00} is the probability that the upper CUSUM equals 0 at time t, given that the upper CUSUM equaled 0 at time $t-1$, that is, $P(S_t^+ = 0 \mid S_{t-1}^+ = 0)$. For the lower one-sided procedure p_{00} is the probability that the lower CUSUM equals 0 at time t, given that the lower CUSUM equaled 0 at time $t-1$, that is, $P(S_t^- = 0 \mid S_{t-1}^- = 0)$. For the two-sided procedure the two one-sided procedures are performed simultaneously. Therefore we have that $p_{00} = P\left(\left\{S_t^+ = 0 \mid S_{t-1}^+ = 0\right\} \cap \left\{S_t^- = 0 \mid S_{t-1}^- = 0\right\}\right)$.

We have that

$$p_{00}$$
$$= P\left(\left\{S_t^+ = 0 \mid S_{t-1}^+ = 0\right\} \cap \left\{S_t^- = 0 \mid S_{t-1}^- = 0\right\}\right)$$

this is computed by substituting in values for h, k, S_t^+, S_{t-1}^+, S_t^- and S_{t-1}^- into equations (2.50) and (2.51)

$$= P\left(\left\{\min\left\{4, \max\left\{0, SN_t - 1 + 0\right\}\right\} = 0\right\} \cap \left\{\max\left\{-4, \min\left\{0, SN_t + 1 + 0\right\}\right\} = 0\right\}\right)$$
$$= P\left(\left\{\max\left\{0, SN_t - 1 + 0\right\} = 0\right\} \cap \left\{\min\left\{0, SN_t + 1 + 0\right\} = 0\right\}\right)$$
$$= P\left(\left(SN_t - 1 \le 0\right) \cap \left(SN_t + 1 + 0 \ge 0\right)\right)$$
$$= P\left(\left(SN_t \le 1\right) \cap \left(SN_t \ge -1\right)\right)$$

recall that $SN_t = 2T - n$ where T is binomially distributed with parameters n and $p = P(X_{ij} \ge \theta_0)$

$$= P\left(\left(2T - 5 \le 1\right) \cap \left(2T - 5 \ge -1\right)\right)$$
$$= P\left(\left(T \le 3\right) \cap \left(T \ge 2\right)\right)$$
$$= P(T = 2) + P(T = 3)$$
$$= 2P(T = 2)$$
$$= 2 \times \left({}^{10}\!/_{32}\right)$$
$$= {}^{20}\!/_{32}$$

since T is binomially distributed with parameters $n = 5$ and $p = 0.5$.

The remaining entries in the first column of the matrix Q can be calculated similarly and we find that $p_{10} = {}^{10}\!/_{32}$ and $p_{20} = {}^{10}\!/_{32}$.

Next we discuss the calculation of the entry in the first row and the second column of the matrix Q, p_{01}, in detail. This is the probability of moving from state 0 to state 1 in one step. This can happen when either the upper cumulative sum or the lower cumulative sum equals 2. But the lower cumulative sum can not equal 2 since by definition the lower cumulative sum can only take on integer values smaller than or equal to zero. Therefore although the required probability is the probability of the *union* of two sets involving values of the upper and the lower CUSUM statistics, one of these sets is empty so that the required probability is the probability of only the non-empty set. Hence, in this case, we will only have to calculate the upper one-sided probability. For the upper one-sided CUSUM, p_{01} is the probability that the upper cumulative sum equals 2 at time t, given that the upper cumulative sum equaled 0 at time $t-1$, that is, $P(S_t^+ = 2 \mid S_{t-1}^+ = 0)$. We have that

$$
\begin{aligned}
& p_{01} \\
&= P(S_t^+ = 2 \mid S_{t-1}^+ = 0) \\
&= P\left(\min\{4, \max\{0, SN_t - 1 + 0\}\} = 2\right) \\
&= P\left(\max\{0, SN_t - 1 + 0\} = 2\right) \\
&= P\left(SN_t - 1 = 2\right) \\
&= P\left(SN_t = 3\right) \\
&= P\left(2T - 5 = 3\right) \\
&= P\left(T = 4\right) \\
&= {}^{5}\!/_{32}.
\end{aligned}
$$

The remaining entries in the second column of the matrix Q can be calculated similarly and we find that $p_{11} = {}^{10}\!/_{32}$ and $p_{21} = {}^{5}\!/_{32}$.

Next we discuss the calculation of the entry in the first row and the third column of the matrix Q, p_{02}, in detail. This is the probability of moving from state 0 to state 2 in one step. This happens when *only* the lower cumulative sum moves to -2, since the upper cumulative sum can not move to -2. Recall that the upper cumulative sum can only take on integer values greater than or equal to zero. As in the case of p_{01}, this probability is also the probability of the *union* of two sets, involving values of the CUSUM statistics, one of which is empty, so that the required probability is the probability of only the non-empty set. Hence, in this case since the lower CUSUM is involved, we will only have to calculate the probability associated with the lower one-sided procedure. Now, for the lower one-sided procedure p_{02} is the probability that the lower cumulative sum equals -2 at time t, given that the lower cumulative sum equaled 0 at time $t-1$, that is, $P(S_t^- = -2 \mid S_{t-1}^- = 0)$. We have that

$$
\begin{aligned}
& p_{02} \\
&= P(S_t^- = -2 \mid S_{t-1}^- = 0) \\
&= P\left(\max\{-4, \min\{0, SN_t + 1 + 0\}\} = -2\right) \\
&= P\left(\min\{0, SN_t + 1 + 0\} = -2\right) \\
&= P\left(SN_t + 1 = -2\right) \\
&= P\left(SN_t = -3\right) \\
&= P\left(2T - 5 = -3\right) \\
&= P\left(T = 1\right) \\
&= {}^{5}\!/_{32}.
\end{aligned}
$$

The remaining entries in the third column of the matrix Q can be calculated similarly and we find that $p_{12} = 5/32$ and $p_{22} = 10/32$.

Next we discuss the calculation of the entry in the first row and the first column of the matrix C, p_{03}, in detail. This is the probability of moving from state 0 to an absorbent state, state 3, in one step. Again, this can happen when *only* the upper cumulative sum moves to 4, since the lower cumulative sum can not move to 4. Recall that the lower cumulative sum can only take on integer values smaller than or equal to zero. Therefore, once again the required probability is the probability of the *union* of two sets involving values of the CUSUM statistics, one of which is empty so that the probability is the probability of only the non-empty set. Therefore we will only have to calculate the upper one-sided probability in this case. For the upper one-sided procedure p_{03} is the probability that the upper cumulative sum equals 4 at time t, given that the upper cumulative sum equaled 0 at time $t-1$, that is, $P(S_t^+ = 4 \mid S_{t-1}^+ = 0)$. We have that

$$
\begin{aligned}
& p_{03} \\
&= P(S_t^+ = 4 \mid S_{t-1}^+ = 0) \\
&= P\left(\min\{4, \max\{0, SN_t - 1 + 0\}\} = 4\right) \\
&= P\left(SN_t - 1 \geq 4\right) \\
&= P\left(SN_t \geq 5\right) \\
&= P\left(2T - 5 \geq 5\right) \\
&= P\left(T \geq 5\right) \\
&= 1/32 .
\end{aligned}
$$

The remaining entries in the first column of the matrix C can be calculated similarly and we find that $p_{13} = 5/32$ and $p_{23} = 1/32$.

Next we discuss the calculation of the entry in the first row and the last column of the matrix C, p_{04}, in detail. This is the probability of moving from state 0 to state 4 in one step. This can happen when *only* the lower cumulative sum moves to -4, since the upper cumulative sum can not move to -4. Recall that the upper cumulative sum can only take on integer values greater than or equal to zero. Therefore although the required probability is the probability of the *union* of two sets involving values of the upper and the lower CUSUM statistics, one of these sets is empty so that the required probability is the probability of only the non-empty

set. Therefore we will only have to calculate the lower one-sided probability. For the lower one-sided procedure p_{04} is the probability that the lower cumulative sum equals -4 at time t, given that the lower cumulative sum equaled 0 at time $t-1$, that is, $P(S_t^- = -4 \mid S_{t-1}^- = 0)$. We have that

$$
\begin{aligned}
& p_{04} \\
&= P(S_t^- = -4 \mid S_{t-1}^- = 0) \\
&= P\left(\max\{-4, \min\{0, SN_t + 1 + 0\}\} = -4\right) \\
&= P\left(SN_t + 1 \le -4\right) \\
&= P\left(SN_t \le -5\right) \\
&= P\left(2T - 5 \le -5\right) \\
&= P\left(T \le 0\right) \\
&= 1/32 .
\end{aligned}
$$

The remaining entries in the last column of the matrix C can be calculated similarly and we find that $p_{14} = 1/32$ and $p_{24} = 6/32$.

The run length distribution and its parameters

The run length distribution and its parameters are calculated using the matrix Q. The ARL is given by $\underline{\xi}(I-Q)^{-1}\underline{1}$ where $\underline{\xi}_{1\times 3} = (1 \quad 0 \quad 0)$, $Q_{3\times 3} = \frac{1}{32}\begin{pmatrix} 20 & 5 & 5 \\ 10 & 10 & 5 \\ 10 & 5 & 10 \end{pmatrix}$ and $\underline{1}_{3\times 1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

As a result, $ARL = E(N) = \underline{\xi}(I-Q)^{-1}\underline{1} = 8.31$.

Let ARL^+ and ARL^- denote the average run lengths of the upper and lower one-sided charts, respectively. The ARL of the two-sided chart can be expressed as a function of the average run lengths of the one-sided charts through the expression

$$ARL = \frac{(ARL^+)(ARL^-)}{\left(ARL^+ + ARL^-\right)} \qquad (2.52)$$

(see Theorem 1 in Appendix A for the proof of result (2.52)).

From the lower- and upper CUSUM sign sections we have that $ARL^{+} = 16.62$ and $ARL^{-} = 16.62$. Using equation (2.52) we have that $ARL = \frac{(16.62)(16.62)}{(16.62+16.62)} = 8.31$.

Table 2.34. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile* values for the two-sided CUSUM sign chart when $n = 5$†.

k	***h***	
	2	**3 or 4**
1	2.67 2.11 (1, 1, 2, 3, 7)	8.31 7.16 (1, 3, 6, 11, 23)
3	16.00 15.50 (1, 5, 11, 22, 47)	‡

Analogous to what was done for the upper one-sided chart, the five percentiles (given in Table 2.34) are displayed in boxplot-like graphs for various h and k values in Figure 2.10. Recall that we would prefer a "boxplot" with a high valued (large) in-control average run length and a small spread. Applying this criterion, we see that the "boxplot" corresponding to the $(h,k) = (2,3)$ combination has the largest in-control average run length, which is favorable, but it also has the largest spread which is unattractive. The "boxplot" furthest to the right is exactly opposite from the "boxplot" furthest to the left. The latter has the smallest spread, which is favorable, but it also has the smallest in-control average run length, which is unattractive. In conclusion, no "box plot" is optimal relative to the others.

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.
† See SAS Program 2 in Appendix B for the calculation of the values in Table 2.34.
‡ Since the decision interval is taken to satisfy $h \le n - k$ there are open cells in Table 2.34.

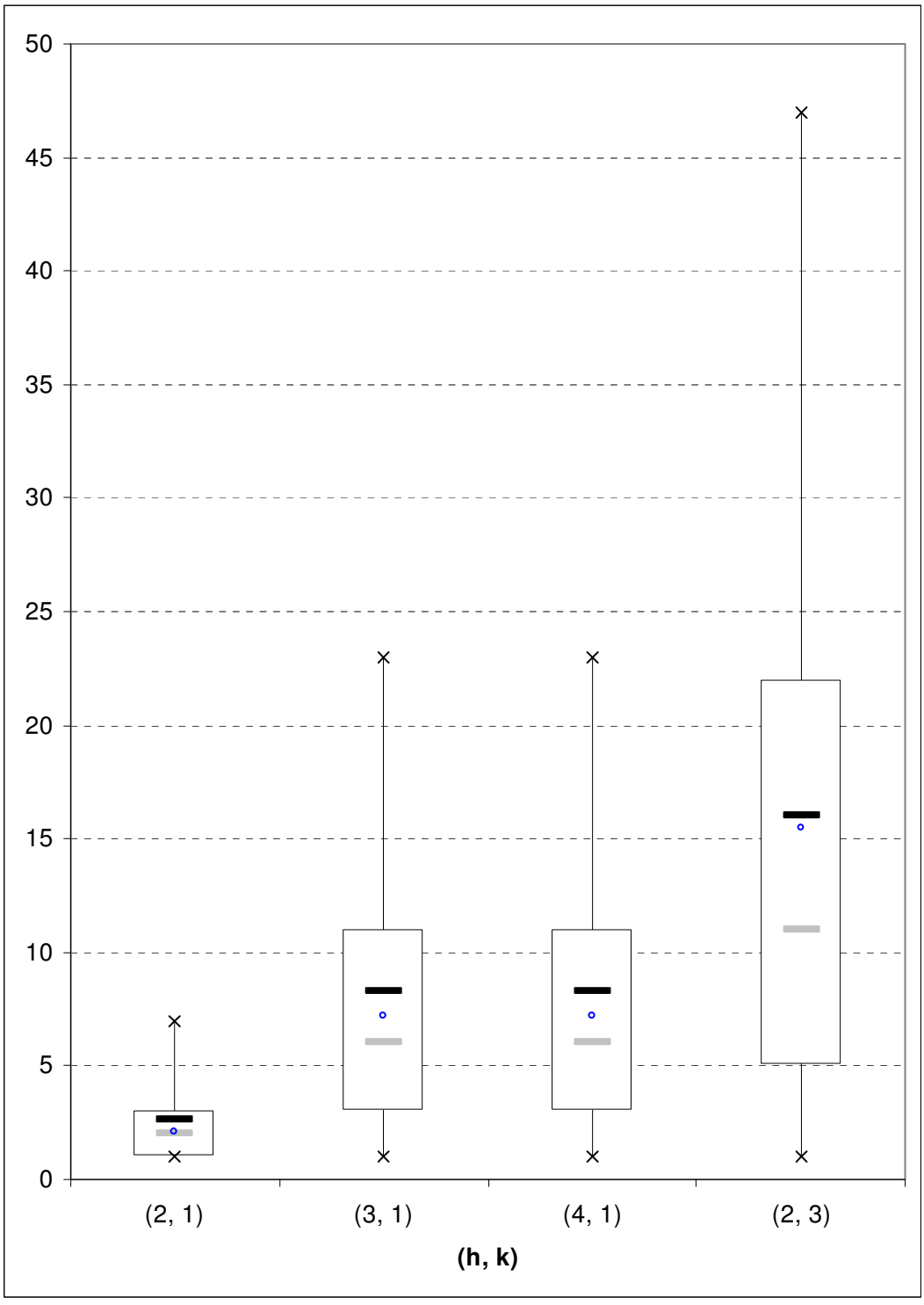

Figure 2.10. Boxplot-like graphs for the in-control run length distribution of various two-sided CUSUM sign charts when $n = 5$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Other values of *h*, *k* and *n* were also considered and the results are given in Table 2.35.

Table 2.35. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile* values for the two-sided CUSUM sign chart when $n = 6$†.

k	***h***		
	2	**3 or 4**	**5 or 6**
0	1.45 0.81 (1, 1, 1, 2, 3)	2.96 1.88 (1, 2, 3, 4, 7)	5.33 4.22 (1, 2, 4, 7, 14)
2	4.57 4.04 (1, 2, 3, 6, 13)	19.34 18.36 (2, 6, 14, 26, 56)	‡
4	32.00 31.50 (2, 10, 22, 44, 95)		

** The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.

† See SAS Program 2 in Appendix B for the calculation of the values in Table 2.35.

‡ Since the decision interval is taken to satisfy $h \le n - k$ there are open cells in Table 2.35.

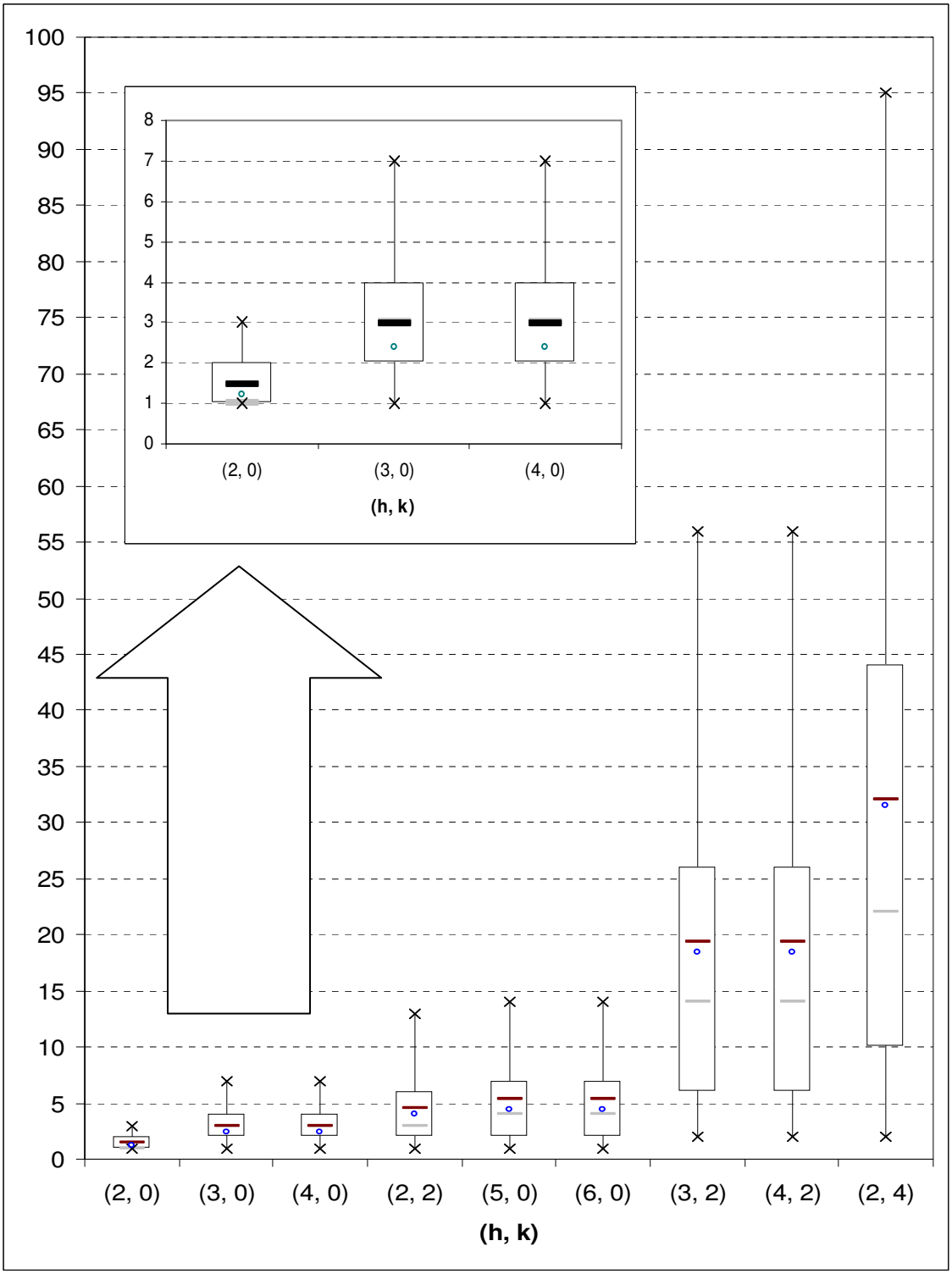

Figure 2.11. Boxplot-like graphs for the in-control run length distribution of various two-sided CUSUM sign charts when $n = 6$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "○" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Table 2.36. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile* values for the two-sided CUSUM sign chart when $n = 10$†.

k	***h***		
	3 or 4	**5 or 6**	**7 or 8**
2	7.17 6.39 (1, 3, 5, 10, 20)	18.41 17.05 (2, 6, 13, 25, 52)	45.80 43.63 (4, 15, 32, 63, 133)
4	38.98 38.30 (3, 12, 27, 54, 115)	232.43 231.26 (13, 68, 161, 322, 694)	‡
6	464.98 464.39 (24, 134, 322, 644, 1392)		

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.

† See SAS Program 2 in Appendix B for the calculation of the values in Table 2.36.

‡ Since the decision interval is taken to satisfy $h \leq n - k$ there are open cells in Table 2.36.

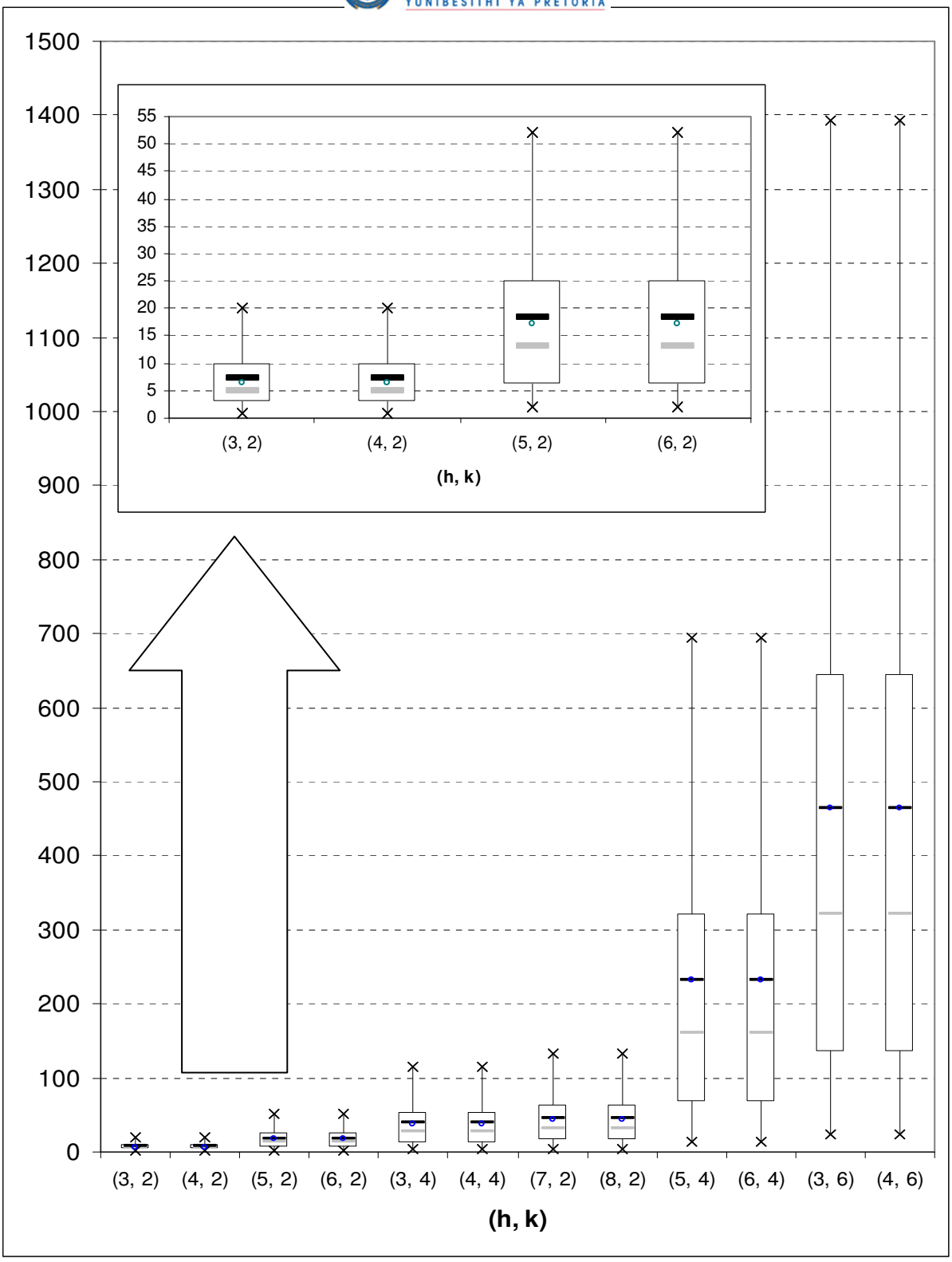

Figure 2.12. Boxplot-like graphs for the in-control run length distribution of various two-sided CUSUM sign charts when $n = 10$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Example 2.13

A two-sided CUSUM sign chart for the Montgomery (2001) piston ring data

We conclude this sub-section by illustrating the two-sided CUSUM sign chart using the piston ring data set from Montgomery (2001). We assume that the underlying distribution is symmetric with a known target value of $\theta_0 = 74\,mm$. Let $k = 3$. Once k is selected, the constant h should be chosen to give the desired in-control average run length performance. By choosing $h = 2$ we obtain an in-control average run length of 16 which is the highest in-control average run length attainable when $n = 5$ (see Table 2.34). Table 2.37 shows the upper and lower sign CUSUM statistics, respectively.

Table 2.37. One-sided sign (S_i^+ and S_i^-) statistics*.

Sample number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
S_i^+	0	0	0	0	0	0	0	0	0	1	0	2	4	6	7
S_i^-	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0

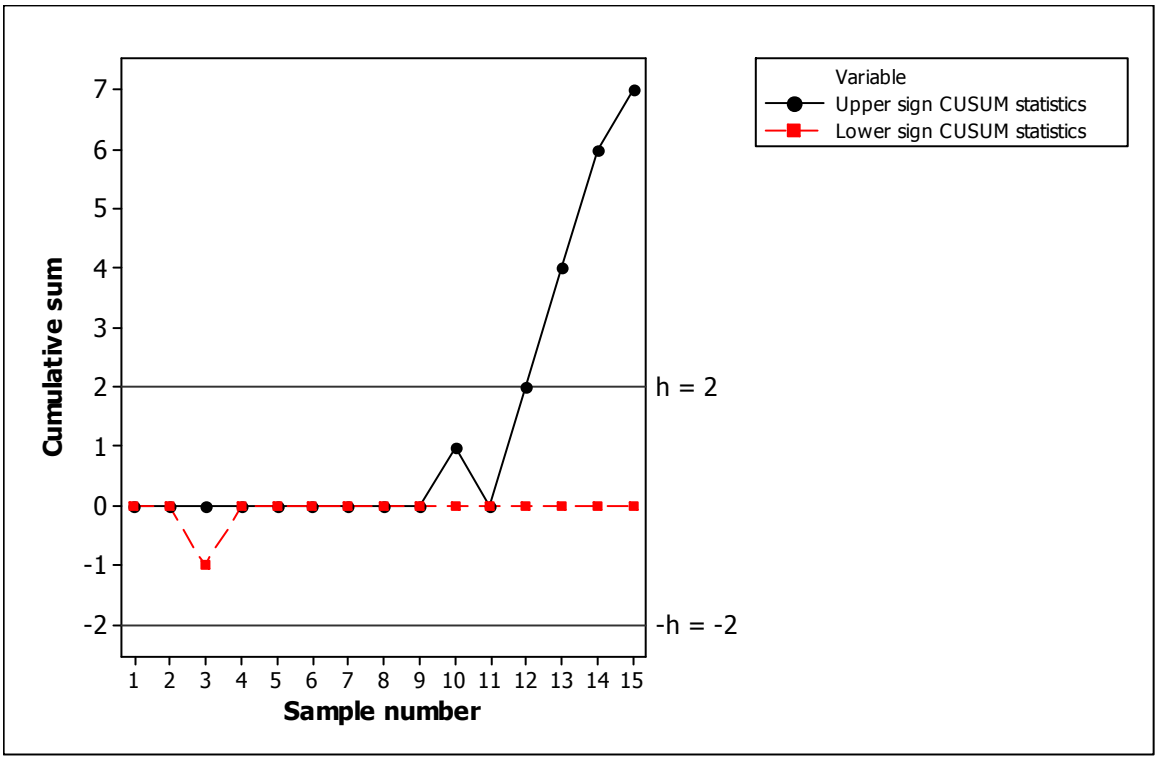

Figure 2.13. The two-sided CUSUM sign chart for the Montgomery (2001) piston ring data.

* See SAS Program 3 in Appendix B for the calculation of the values in Table 2.37.

The two-sided CUSUM sign chart signals at sample number 12, indicating a most like upward shift in the process median. The action taken following an out-of-control signal on a CUSUM chart is identical to that with any control chart. A search for assignable causes should be done, corrective action should be taken (if required) and, following this, the CUSUM is reset to zero.

2.3.4. Summary

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. In this section we have described the properties of the CUSUM sign chart and given tables for its implementation. Detailed calculations have been given to help the reader to understand the subject more thoroughly.

2.4. The EWMA control chart

2.4.1. Introduction

The exponentially weighted moving average (EWMA) scheme was first introduced by Roberts (1959). In a subsequent article, Roberts (1966) compared the performance of EWMA charts to Shewhart and CUSUM charts. Various authors have studied EWMA charts (see for example Robinson and Ho (1978) and Crowder (1987)). EWMA charts have become very popular over the last few years. It is one of several charting methods aimed at correcting a deficiency of the Shewhart chart – insensitivity to small process shifts.

An EWMA control chart scheme accumulates statistics $X_1, X_2, X_3, \ldots$ with the plotting statistics defined as

$$Z_i = \lambda X_i + (1-\lambda) Z_{i-1} \tag{2.53}$$

where $0 < \lambda \leq 1$ is a constant called the weighting constant. The starting value Z_0 is often taken to be the process target value, i.e. $Z_0 = \theta_0$.

The EWMA chart is constructed by plotting Z_i against the sample number i (or time). If the plotting statistic Z_i falls between the two control limits, that is, $LCL < Z_i < UCL$, the process is considered to be in-control. If the plotting 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 out-of-control.

To illustrate that the plotting statistic Z_i is a weighted average of all the previous statistics, Z_{i-1} may be substituted by $Z_{i-1} = \lambda X_{i-1} + (1-\lambda)Z_{i-2}$ in equation (2.53) to obtain

$$\begin{aligned} Z_i &= \lambda X_i + (1-\lambda)\left(\lambda X_{i-1} + (1-\lambda)Z_{i-2}\right) \\ &= \lambda X_i + \lambda(1-\lambda)X_{i-1} + (1-\lambda)^2 Z_{i-2} \\ &= \lambda X_i + \lambda(1-\lambda)X_{i-1} + (1-\lambda)^2\left(\lambda X_{i-2} + (1-\lambda)Z_{i-3}\right) \\ &= \lambda X_i + \lambda(1-\lambda)X_{i-1} + \lambda(1-\lambda)^2 X_{i-2} + (1-\lambda)^3 Z_{i-3}. \end{aligned}$$

This method of substitution is called *recursive substitution.* By continuing the process of recursive substitution for Z_{i-p}, $p = 2,3,...,t$, we obtain

$$Z_i = \lambda\sum_{p=0}^{i-1}(1-\lambda)^p X_{i-p} + (1-\lambda)^i Z_0. \qquad (2.54)$$

We can see from expression (2.54) that Z_i can be written as a moving average of the current and past observations which has geometrically decreasing weights $\lambda(1-\lambda)^p$ associated with increasingly aged observations X_{i-p} ($p = 1,2,...$). Therefore, the EWMA has been referred to as a geometric moving average (see, for example, Montgomery (2005)).

If the observations $\{X_i, i = 1,2,...\}$ are independent identically distributed variables with mean μ and variance σ^2, then the mean and the variance of the plotting statistic Z_i are given by

$$\mu_{Z_i} = E(Z_i) = \mu \quad \text{for} \quad i = 1,2,...$$

and

$$\sigma^2_{Z_i} = \sigma^2\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right) \quad \text{for} \quad i = 1,2,... \ .$$

The *exact* control limits and the center line of the EWMA control chart are given by

$$\begin{aligned} UCL &= \theta_0 + L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \\ CL &= \theta_0 \\ LCL &= \theta_0 - L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \end{aligned} . \qquad (2.55)$$

From (2.55) we see that we have two design parameters of interest, namely the multiplier L, ($L > 0$) and the smoothing constant λ. We also see that $\left(1-(1-\lambda)^{2i}\right) \to 1$ as i increases. Therefore, as i increases the control limits will approach steady-state values given by

$$\begin{aligned} UCL &= \theta_0 + L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \\ LCL &= \theta_0 - L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \end{aligned} . \qquad (2.56)$$

The above-mentioned control limits are called *steady-state* control limits.

Various authors recommend choosing the EWMA constants L and λ by minimizing the average run length at a specified shift for a desired in-control average run length. In general, values of λ in the interval $0.05 \le \lambda \le 0.25$ work well in the normal theory case with $\lambda = 0.05$, $\lambda = 0.1$ and $\lambda = 0.2$ being popular choices. The ARL_0, standard deviation of the run length ($SDRL$), 5^{th}, 25^{th} (the first quartile, Q_1), 50^{th} (the median run length, MRL), 75^{th} (the third quartile, Q_3) and 95^{th} percentile values can be computed and tabulated for various values of L and λ.

Lucas and Saccucci (1990) have investigated some properties of the EWMA chart under the assumption of independent normally distributed observations. Lucas and Saccucci's most important contribution is the use of a Markov-chain approach to evaluate the run-length properties of the EWMA chart. It is important to note that the successive observations are assumed to be independent over time in their evaluation. Lucas and Saccucci (1990) used a procedure similar to that described by Brook and Evans (1972) to approximate the properties of an EWMA scheme. They evaluate the properties of the *continuous state* Markov chain by

discretizing the infinite state transition probability matrix (TPM). This procedure entails dividing the interval between the upper control limit and the lower control limit into N subintervals of width 2δ. Then the plotting statistic, Z_i, is said to be in the non-absorbing state j at time i if

$$S_j - \delta < Z_i \leq S_j + \delta \quad \text{for} \quad j = 1,2,...,N-1$$

and

$$S_j - \delta < Z_i < S_j + \delta \quad \text{for} \quad j = N$$

where S_j denotes the midpoint of the j^{th} interval. Let r denote the number of non-absorbing states. Z_i is said to be in the absorbing state if Z_i falls on or outside one of the control limits, that is, $Z_i \leq LCL$ or $Z_i \geq UCL$. Clearly, there are $r+1$ states, since there are r non-absorbing states and one absorbing state. Lucas and Saccucci (1990) have done a thorough job of evaluating the run length properties of the EWMA chart and provided helpful tables for the design of the EWMA chart. Additional tables are provided in the technical report by Lucas and Saccucci (1987). In their 1990 paper they concentrate on the average run length characteristics of various charting combinations. The authors conclude that EWMA procedures have average run length properties similar to those for CUSUM procedures. This point has also been made by various authors, for example, Ewan (1963), Roberts (1966) and Montgomery, Gardiner and Pizzano (1987). In this thesis, the approach taken by Lucas and Saccucci (1990) is extended to the use of the sign statistic resulting in an EWMA sign chart that accumulates the statistics $SN_1, SN_2, SN_3, \ldots$.

2.4.2. The proposed EWMA sign chart

A nonparametric EWMA-type of control chart based on the sign statistic can be obtained by replacing X_i in expression (2.53) with SN_i (recall that $SN_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0)$).

The EWMA sign chart accumulates the statistics $SN_1, SN_2, SN_3, \ldots$ with the plotting statistics defined as

$$Z_i = \lambda SN_i + (1-\lambda)Z_{i-1} \tag{2.57}$$

where $0 < \lambda \leq 1$ and the starting value Z_0 is usually taken to equal zero, i.e. $Z_0 = 0$.

The expected value, variance and standard deviation of SN_i are found from the fact that the distribution of SN_i can easily be obtained from the distribution of the binomial distribution (recall that $SN_i = 2T_i - n$ if there are no ties within a subgroup, where T_i has a binomial distribution with parameters n and $p = P(X_{ij} \geq \theta_0))$. The formulas for the expected value, variance and standard deviation of SN_i was derived in Section 2.1 and we obtained $E(SN_i) = n(2p-1)$, $\text{var}(SN_i) = 4np(1-p)$ and $stdev(SN_i) = \sigma_{SN_i} = 2\sqrt{np(1-p)}$, respectively. The starting value Z_0 can also be taken to be the expected value of SN_i, therefore $Z_0 = E(SN_i) = n(2p-1)$ and in the in-control case where $p = 0.5$ we have $Z_0 = n(2\times 0.5 - 1) = 0$ for all n.

From the similarity between the definitions of the normal EWMA scheme and the sign EWMA scheme, it follows that the exact control limits and the center line of the EWMA sign control chart can be obtained by replacing σ in (2.55) with σ_{SN_i} which yields

$$\begin{aligned} UCL &= 0 + L\sigma_{SN_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \\ CL &= 0 \\ LCL &= 0 - L\sigma_{SN_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \end{aligned} . \qquad (2.58)$$

It is important to note θ_0 in (2.55) is replaced by 0 in (2.58). This is because the EWMA sign chart is designed for the sign test statistic and not for the observations (the X_i's).

Similarly, the *steady-state* control limits can be obtained by replacing σ in (2.56) with σ_{SN_i} and θ_0 by zero which yields

$$\begin{aligned} UCL &= L\sigma_{SN_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \\ LCL &= -L\sigma_{SN_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \end{aligned} . \qquad (2.59)$$

2.4.3. Markov-chain approach

Lucas and Saccucci (1990) evaluated the properties of the *continuous state* Markov chain by *discretizing* the infinite state TPM. This procedure entails dividing the interval between the UCL and the LCL into N subintervals of width 2δ. The width of each subinterval can be obtained by setting $\delta = \frac{1}{2} \times \frac{(UCL - LCL)}{N}$ and we get that the width of an interval equals $2\delta = \frac{(UCL - LCL)}{N}$. Thus, the endpoints of the subintervals will be given by $LCL,\ \ LCL + \frac{(UCL - LCL)}{N},\ LCL + 2\frac{(UCL - LCL)}{N}, \ldots,\ LCL + (N-1)\frac{(UCL - LCL)}{N},\ UCL,$ respectively (see Figure 2.14). In general, the endpoints of the j^{th} interval will be given by

$$(LCL_j, UCL_j) = \left(LCL + (j-1) \times \frac{(UCL - LCL)}{N}, LCL + j \times \frac{(UCL - LCL)}{N} \right).$$

The midpoint of the j^{th} interval, S_j, is easily obtained by taking the sum of the two endpoints of the j^{th} interval and dividing it by 2. Thus, we obtain

$$\begin{aligned} & S_j \\ &= \frac{LCL_j + UCL_j}{2} \\ &= \frac{\left(LCL + \frac{(j-1)(UCL - LCL)}{N} \right) + \left(LCL + \frac{j(UCL - LCL)}{N} \right)}{2} \\ &= \frac{(2j-1)(UCL - LCL)}{2N}. \end{aligned}$$

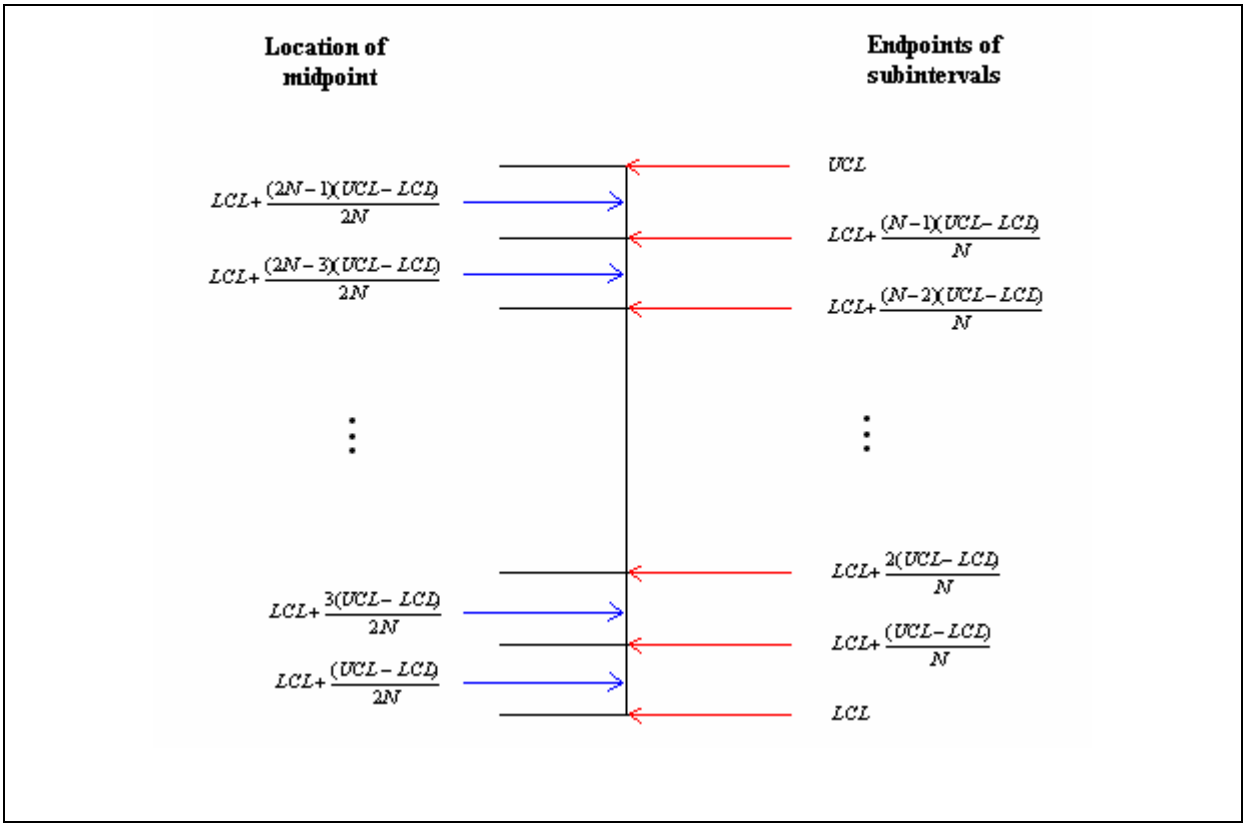

Figure 2.14. Partitioning of the interval between the UCL and the LCL into N subintervals.

Then the plotting statistic, Z_i, is said to be in the non-absorbing state j at time i if

$$S_j - \delta < Z_i \leq S_j + \delta \quad \text{for} \quad j = 1,2,...,N-1 \tag{2.60}$$

and

$$S_j - \delta < Z_i < S_j + \delta \quad \text{for} \quad j = N\,. \tag{2.61}$$

Z_i is said to be in the absorbing state if Z_i falls on or outside one of the control limits, that is, $Z_i \leq LCL$ or $Z_i \geq UCL$.

Let p_{ij} denote the probability of moving from state i to state j in one step. We have that $p_{ij} = P(\text{Moving to state } j \mid \text{in state } i)$. To calculate this probability we assume that the plotting statistic is equal to S_i whenever it is in state i. For all j non-absorbing we obtain $p_{ij} = P(S_j - \delta < Z_k \leq S_j + \delta \mid Z_{k-1} = S_i)$. This is the probability that Z_k is within state j, conditioned on Z_{k-1} being equal to the midpoint of state i. By using the definition of the plotting statistic given in expression (2.57) this transition probability can be written as

$$\begin{aligned} p_{ij} &= P\left(S_j - \delta < \lambda SN_k + (1-\lambda)Z_{k-1} \le S_j + \delta \mid Z_{k-1} = S_i\right) \\ &= P\left(S_j - \delta < \lambda SN_k + (1-\lambda)S_i \le S_j + \delta\right) \\ &= P\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} < SN_k \le \frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda}\right). \end{aligned}$$

Recall that $SN_k = 2T_k - n$ where T_k is binomially distributed with parameters n and $p = P(X_{ij} \ge \theta_0)$

$$\begin{aligned} p_{ij} &= P\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} < 2T_k - n \le \frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(\frac{\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} + n\right)}{2} < T_k \le \frac{\left(\frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda} + n\right)}{2}\right). \end{aligned} \tag{2.62}$$

The probability of transition to the out-of-control state can be determined similarly. For all j absorbing we obtain

$$\begin{aligned} p_{ij} &= P\left(Z_k \le LCL \mid Z_{k-1} = S_i\right) + P\left(Z_k \ge UCL \mid Z_{k-1} = S_i\right) \\ &= P\left(\lambda SN_k + (1-\lambda)Z_{k-1} \le LCL \mid Z_{k-1} = S_i\right) + P\left(\lambda SN_k + (1-\lambda)Z_{k-1} \ge UCL \mid Z_{k-1} = S_i\right) \\ &= P\left(\lambda SN_k + (1-\lambda)S_i \le LCL\right) + P\left(\lambda SN_k + (1-\lambda)S_i \ge UCL\right) \\ &= P\left(SN_k \le \frac{LCL - (1-\lambda)S_i}{\lambda}\right) + P\left(SN_k \ge \frac{UCL - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(2T_k - n \le \frac{LCL - (1-\lambda)S_i}{\lambda}\right) + P\left(2T_k - n \ge \frac{UCL - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(T_k \le \frac{\left(\frac{LCL - (1-\lambda)S_i}{\lambda} + n\right)}{2}\right) + P\left(T_k \ge \frac{\left(\frac{UCL - (1-\lambda)S_i}{\lambda} + n\right)}{2}\right). \end{aligned} \tag{2.63}$$

Since the values δ, λ, n, S_i and S_j are known constants the binomial probabilities in expressions (2.62) and (2.63) can easily be calculated using some type of statistical software package, for example, Excel or SAS.

Once the one-step transition probabilities are calculated, the TPM can be constructed and is given by $TPM = \begin{pmatrix} Q & | & \underline{p} \\ - & - & - \\ \underline{0}' & | & 1 \end{pmatrix}$ (written in partitioned form) where Q is the matrix that contains all the transition probabilities of going from a non-absorbing state to a non-absorbing state. In other words, Q is the transition matrix among the in-control states, $Q:(NA \to NA)$. $\underline{p}$ contains all the transition probabilities of going from each non-absorbing state to the absorbing states, $\underline{p}:(NA \to A)$. $\underline{0}' = (0 \quad 0 \quad 0 \quad \cdots \quad 0)$ contains all the transition probabilities of going from the absorbing state to each non-absorbing state, $\underline{0}' :(A \to NA)$. 1 represents the scalar value one which is the probability of going from the absorbing state to the absorbing state, $1:(A \to A)$.

Lucas and Saccucci (1990) have investigated some properties of the EWMA chart under the assumptions of independent normally distributed observations. From the similarity between the definitions of the normal EWMA scheme and the sign EWMA scheme, it follows that the formulas derived by Lucas and Saccucci (1990) can be extended to the use of the sign EWMA scheme. The formulas derived by Lucas and Saccucci (1990) have been studied by other authors, for example, Fu, Spiring and Xie (2002) and Fu and Lou (2003). The latter two used the moment generating function and the probability generating function, respectively, to derive expressions for the first and second moments of the run length variable N. See Theorem 2 in Appendix A for the derivations done by Fu, Spiring and Xie (2002) and Fu and Lou (2003). For the formulas refer to equations (2.41) to (2.45) of this thesis.

Example 2.14

The EWMA sign chart where the sample size is even (*n* = 6)

We consider the EWMA sign chart with a smoothing constant of 0.1 ($\lambda = 0.1$) and a multiplier of 3 ($L = 3$). The interval between the UCL and the LCL is divided into 4 subintervals ($N = 4$). For a sample size of 6, the sign statistic SN_i can take on the values $\{-6,-4,-2,0,2,4,6\}$ and the statistic T_i takes on the values $\{0,1,2,3,4,5,6\}$.

The *steady-state* control limits are given in (2.59) by

$$UCL = L\sigma_{SN_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

$$LCL = -L\sigma_{SN_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

where $L = 3$, $\lambda = 0.1$, and $\sigma_{SN_i} = 2.449$, since $\sigma_{SN_i} = 2\sqrt{np(1-p)} = 2\sqrt{6(0.5)(1-0.5)} = 2.449$.

Clearly, we only have to calculate the UCL since $LCL = -UCL$. We have that $UCL = 3\times 2.449\sqrt{\left(\frac{0.1}{2-0.1}\right)} = 1.686$. Therefore, $LCL = -1.686$.

This Markov-chain procedure entails dividing the interval between the UCL and the LCL into subintervals of width 2δ. Figure 2.15 illustrates the partitioning of the interval between the UCL and the LCL into subintervals.

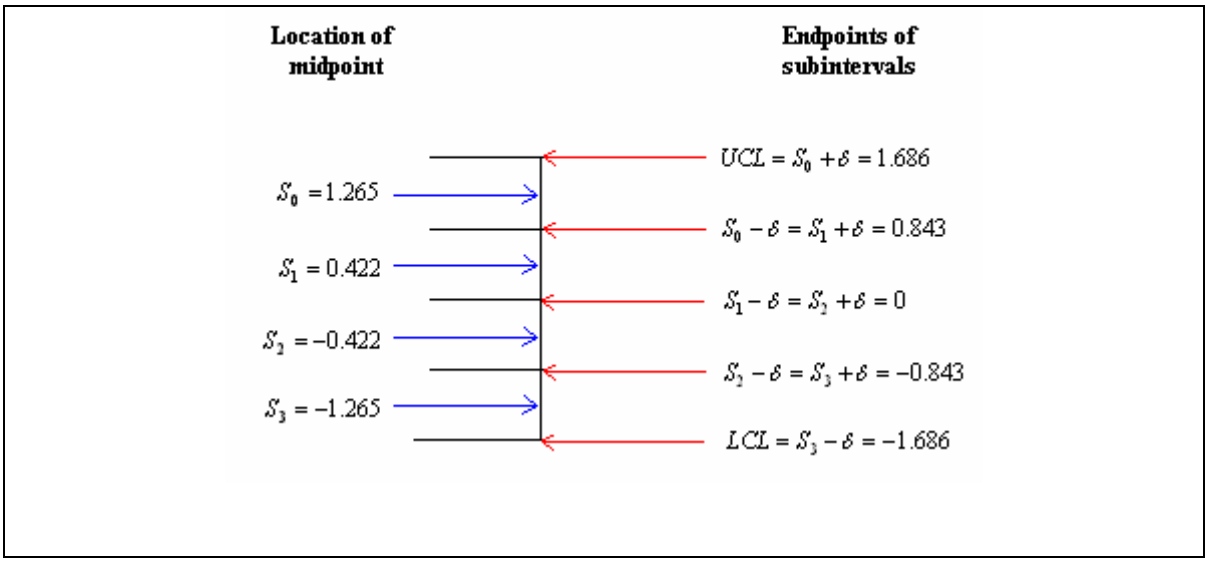

Figure 2.15. Partitioning of the interval between the UCL and the LCL into 4 subintervals.

From Figure 2.15 we see that there are 4 non-absorbing states, i.e. $r = 4$. The TPM is given by

$$TPM_{5\times5} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \begin{pmatrix} Q_{4\times4} & | & \underline{p}_{4\times1} \\ - & - & - \\ \underline{0}'_{1\times4} & | & 1_{1\times1} \end{pmatrix}.$$

The plotting statistic, Z_i, is said to be in the non-absorbing state j at time i if $S_j - \delta < Z_i \leq S_j + \delta$ for $j = 0,1,2,3$ where S_j denotes the midpoint of the j^{th} interval. Each sub-interval has a width of $2\delta = 0.843$, therefore $\delta = 0.4215$.

Table 2.38. Calculation of the one-step probabilities in the first row of the TPM.

$p_{00} = P(\textit{Moving to state } 0 \mid \textit{in state } 0)$

$= P(S_0 - \delta < Z_k \le S_0 + \delta \mid Z_{k-1} = S_0)$

using expression (2.62) we obtain

$$= P\left(\frac{\left(\frac{(S_0 - \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} < T_k \le \frac{\left(\frac{(S_0 + \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} \right)$$

with $\delta = 0.4215$, $\lambda = 0.1$, $L = 3$ and $S_0 = 1.265$

$= P(1.525 < T_k \le 5.740)$

$= P(T_k = 2) + P(T_k = 3) + P(T_k = 4) + P(T_k = 5)$

$= \frac{56}{64}$

$p_{01} = P(\textit{Moving to state } 1 \mid \textit{in state } 0)$

$= P(S_1 - \delta < Z_k \le S_1 + \delta \mid Z_{k-1} = S_0)$

using expression (2.62) we obtain

$$= P\left(\frac{\left(\frac{(S_1 - \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} < T_k \le \frac{\left(\frac{(S_1 + \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} \right)$$

$= P(-2.690 < T_k \le 1.525)$

$= P(T_k \le 1)$

$= \frac{7}{64}$

$$p_{02} = P(\textit{Moving to state } 2 \mid \textit{in state } 0)$$

$$= P(S_2 - \delta < Z_k \le S_2 + \delta \mid Z_{k-1} = S_0)$$

using expression (2.62) we obtain

$$= P\left(\frac{\left(\frac{(S_2 - \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} < T_k \le \frac{\left(\frac{(S_2 + \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} \right)$$

$$= P(-6.904 < T_k \le -2.690)$$

$$= 0$$

$$p_{03} = P(\textit{Moving to state } 3 \mid \textit{in state } 0)$$

$$= P(S_3 - \delta < Z_k \le S_3 + \delta \mid Z_{k-1} = S_0)$$

using expression (2.62) we obtain

$$= P\left(\frac{\left(\frac{(S_3 - \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} < T_k \le \frac{\left(\frac{(S_3 + \delta) - (1-\lambda)S_0}{\lambda} + n \right)}{2} \right)$$

$$= P(-11.119 < T_k \le -6.904)$$

$$= 0$$

$$p_{04} = P(\textit{Moving to state } 4 \mid \textit{in state } 0)$$

$$= P(Z_k \le LCL \mid Z_{k-1} = S_0) + P(Z_k \ge UCL \mid Z_{k-1} = S_0)$$

using expression (2.63) we obtain

$$= P\left(T_k \le \frac{LCL - (1-\lambda)S_0}{\lambda} + n \right) + P\left(T_k \ge \frac{UCL - (1-\lambda)S_0}{\lambda} + n \right)$$

$$= P(T_k \le -11.119) + P(T_k \ge 5.739)$$

$$= 0 + P(T_k = 6)$$

$$= \frac{1}{64}$$

The calculations of the other transition probabilities can be done similarly. Therefore the TPM is given by $TPM = \begin{pmatrix} 56/64 & 7/64 & 0 & 0 & 1/64 \\ 1/64 & 56/64 & 7/64 & 0 & 0 \\ 0 & 7/64 & 56/64 & 1/64 & 0 \\ 0 & 0 & 7/64 & 56/64 & 1/64 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_{4\times4} & | & \underline{p}_{4\times1} \\ - & - & - \\ \underline{0}'_{1\times4} & | & 1_{1\times1} \end{pmatrix}$.

Other values of the multiplier (L) and the smoothing constant (λ) were also considered and the results are given in Tables 2.39 and 2.40*.

Table 2.39. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile values† for the EWMA sign chart when $n = 6$ and $N = 5$, i.e. there are 5 subintervals between the lower and upper control limit‡.

	$L = 1$	$L = 2$	$L = 3$
$\lambda = 0.05$	6.79 8.56 (1, 1, 3, 9, 24)	22.86 30.29 (1, 3, 10, 31, 85)	736.00 827.24 (4, 134, 472, 1051, 2393)
$\lambda = 0.1$	4.84 5.36 (1, 1, 3, 7, 16)	83.69 104.13 (1, 6, 47, 121, 294)	736.00 819.78 (4, 142, 477, 1049, 2377)
$\lambda = 0.2$	4.73 5.08 (1, 1, 3, 6, 15)	34.12 39.63 (1, 5, 21, 49, 114)	585.80 608.31 (9, 152, 398, 820, 1800)

Similar tables can be constructed by changing the sample size (n), the number of subintervals between the lower and upper control limit (N), the multiplier (L) and the smoothing constant (λ) in the SAS program for the EWMA sign chart given in Appendix B.

* These results were calculated through the formulas given in equations (2.41) to (2.45).
† The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.
‡ See SAS Program 4 in Appendix B for the calculation of the values in Table 2.39.

Table 2.40. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for the EWMA sign chart when $n = 10$ and $N = 5$, i.e. there are 5 subintervals between the lower and upper control limit†.

	$L = 1$	$L = 2$	$L = 3$
$\lambda = 0.05$	25.47 31.96 (1, 2, 13, 37, 90)	166.06 207.27 (1, 11, 93, 241, 585)	1773.34 2089.12 (6, 228, 1087, 2557, 5970)
$\lambda = 0.1$	10.35 11.02 (1, 2, 7, 14, 33)	75.61 81.91 (1, 16, 50, 107, 204)	845.42 890.53 (9, 208, 570, 1188, 2624)
$\lambda = 0.2$	3.67 3.68 (1, 1, 2, 5, 11)	25.47 31.96 (1, 2, 13, 37, 90)	272.79 305.97 (1, 51, 176, 389, 886)

From Tables 2.39 and 2.40 we see that the ARL_0, $SDRL$ and percentiles increase as the value of the multiplier (L) increases. In contrast, the ARL_0, $SDRL$ and percentiles decrease as the value of the smoothing constant (λ) increases. From Table 2.40 we find an in-control average run length of 272.79 for $n = 10$ when the multiplier is taken to equal 3 ($L = 3$) and the smoothing constant 0.2 ($\lambda = 0.2$). The chart performance is good, since the attained in-control average run length of 272.79 is in the region of the desired in-control average run length which is generally taken to be 370 or 500.

2.4.4. Summary

EWMA charts are popular control charts; they 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 implement than the CUSUM charts. We have described the properties of the EWMA sign chart and given tables for its implementation. Although a lot has been done over the past few years concerning EWMA-type charts, more work is necessary on the practical implementation of the charts as well as on adaptations in case U.

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.
† See SAS Program 4 in Appendix B for the calculation of the values in Table 2.40.

Chapter 3: Signed-rank charts

3.1. The Shewhart-type control chart

3.1.1. Introduction

As mentioned in Chapter 2, samples of fixed size are taken at regular intervals and the plotting statistic is then plotted. The question is: Which quality parameter should be used as the plotting statistic? In Chapter 2 the sign test statistic SN_i was described and it was mentioned that the sign test statistic is only influenced by the signs of the deviations $(x_{ij} - \theta_0)$. There is an alternative statistic that can be used to track the location of a process. The statistic is a function of both the magnitudes and signs of the $(x_{ij} - \theta_0)$'s, called the signed-rank statistic.

3.1.2. Definition of the signed-rank test statistic

The signed-rank test is a nonparametric test that can be used to test hypotheses on or construct confidence intervals (see Gibbons and Chakraborti (2003)) for the median of any symmetric continuous population distribution. Let $X_{i1}, X_{i2}, ..., X_{in}$ denote the i^{th} $(i = 1,2,...)$ sample or subgroup of independent observations of size $n > 1$ from a process with an unknown continuous distribution function denoted by F. Let θ_0 denote the known in-control location parameter (also called the target value). Let R_{ij}^{+} denote the rank of the absolute deviations, $|x_{ij} - \theta_0|$, within the subgroup $(|x_{i1} - \theta_0|, |x_{i2} - \theta_0|, ..., |x_{in} - \theta_0|)$ for $i = 1,2,3...$. Then R_{ij}^{+} is referred to as the within-group absolute rank of the deviations. The signed-rank test statistic is given by

$$SR_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0) R_{ij}^{+} \text{ for } i = 1,2,3... \tag{3.1}$$

where $sign(x) = -1, 0, 1$ if $x < 0, = 0, > 0$.

3.1.3. Plotting statistic

The signed-rank test statistic, SR_i (given in (3.1)), is used as the plotting statistic on the Shewhart signed-rank chart. If the plotting statistic SR_i falls between the two control limits, that is, $LCL < SR_i < UCL$, the process is considered to be in-control. If the plotting statistic SR_i falls on or outside one of the control limits, that is $SR_i \le LCL$ or $SR_i \ge UCL$, the process is considered to be out-of-control.

The plotting statistic is linearly related to the well-known Wilcoxon signed-rank statistic T_n^+ through the formula (see Bakir (2003), page 424, equation 2.4)

$$SR_i = 2T_n^+ - \frac{n(n+1)}{2} \tag{3.2}$$

where $T_n^+ = \sum_{j=1}^{n} \psi(x_{ij} - \theta_0) R_{ij}^+$, $\psi(x) = 0, 1$ if $x \le 0, > 0$.

Example 3.1

A two-sided Shewhart signed-rank chart for the Montgomery (2001) piston ring data

We illustrate the Shewhart-type signed-rank chart using the same set of data from Montgomery (2001) that was used in example 2.1. We assume that the underlying distribution is symmetric with a known median $\theta_0 = 74\,mm$. Panel *a* of Table 3.1 exhibits the individual observations of 15 independent samples, each of size 5 i.e. $n = 5$. The absolute deviations $|x_{ij} - \theta_0|$ and $sign(x_{ij} - \theta_0)$ are shown in panel *b* and panel *c* of Table 3.1, respectively. The known target value is taken to be 74, that is, $\theta_0 = 74$. The within-group absolute rank of the deviations R_{ij}^+ and the $sign(x_{ij} - \theta_0) R_{ij}^+$ values are shown in panel *a* and panel *b* of Table 3.2, respectively. Panel *c* of Table 3.2 holds the signed-ranks i.e. SR_i for $i = 1,2,3,...,15$.

Table 3.1. Data and calculations for the signed-rank chart*.

Sample number	Panel *a*					Panel *b*					Panel *c*				
	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}	$\lvert x_{i1}-\theta_0\rvert$	$\lvert x_{i2}-\theta_0\rvert$	$\lvert x_{i3}-\theta_0\rvert$	$\lvert x_{i4}-\theta_0\rvert$	$\lvert x_{i5}-\theta_0\rvert$	$sign(x_{i1}-\theta_0)$	$sign(x_{i2}-\theta_0)$	$sign(x_{i3}-\theta_0)$	$sign(x_{i4}-\theta_0)$	$sign(x_{i5}-\theta_0)$
1	74.012	74.015	74.030	73.986	74.000	0.012	0.015	0.030	0.014	0.000	1	1	1	-1	0
2	73.995	74.010	73.990	74.015	74.001	0.005	0.010	0.010	0.015	0.001	-1	1	-1	1	1
3	73.987	73.999	73.985	74.000	73.990	0.013	0.001	0.015	0.000	0.010	-1	-1	-1	0	-1
4	74.008	74.010	74.003	73.991	74.006	0.008	0.010	0.003	0.009	0.006	1	1	1	-1	1
5	74.003	74.000	74.001	73.986	73.997	0.003	0.000	0.001	0.014	0.003	1	0	1	-1	-1
6	73.994	74.003	74.015	74.020	74.004	0.006	0.003	0.015	0.020	0.004	-1	1	1	1	1
7	74.008	74.002	74.018	73.995	74.005	0.008	0.002	0.018	0.005	0.005	1	1	1	-1	1
8	74.001	74.004	73.990	73.996	73.998	0.001	0.004	0.010	0.004	0.002	1	1	-1	-1	-1
9	74.015	74.000	74.016	74.025	74.000	0.015	0.000	0.016	0.025	0.000	1	0	1	1	0
10	74.030	74.005	74.000	74.016	74.012	0.030	0.005	0.000	0.016	0.012	1	1	0	1	1
11	74.001	73.990	73.995	74.010	74.024	0.001	0.010	0.005	0.010	0.024	1	-1	-1	1	1
12	74.015	74.020	74.024	74.005	74.019	0.015	0.020	0.024	0.005	0.019	1	1	1	1	1
13	74.035	74.010	74.012	74.015	74.026	0.035	0.010	0.012	0.015	0.026	1	1	1	1	1
14	74.017	74.013	74.036	74.025	74.026	0.017	0.013	0.036	0.025	0.026	1	1	1	1	1
15	74.010	74.005	74.029	74.000	74.020	0.010	0.005	0.029	0.000	0.020	1	1	1	0	1

* See SAS Program 5 in Appendix B for the calculation of the values in Table 3.1.

Table 3.2. Calculations for the signed-rank chart*.

Sample number	Panel *a*					Panel *b*					Panel *c*
	R_{i1}^+	R_{i2}^+	R_{i3}^+	R_{i4}^+	R_{i5}^+	$sign(x_{i1}-\theta_0)R_{i1}^+$	$sign(x_{i2}-\theta_0)R_{i2}^+$	$sign(x_{i3}-\theta_0)R_{i3}^+$	$sign(x_{i4}-\theta_0)R_{i4}^+$	$sign(x_{i5}-\theta_0)R_{i5}^+$	SR_i
1	2	4	5	3	1	2	4	5	-3	0	8
2	2	4	4	5	1	-2	4	-4	5	1	4
3	4	2	5	1	3	-4	-2	-5	0	-3	-14
4	3	5	1	4	2	3	5	1	-4	2	7
5	4	1	2	5	4	4	0	2	-5	-4	-3
6	3	1	4	5	2	-3	1	4	5	2	9
7	4	1	5	3	3	4	1	5	-3	3	10
8	1	4	5	4	2	1	4	-5	-4	-2	-6
9	3	2	4	5	2	3	0	4	5	0	12
10	5	2	1	4	3	5	2	0	4	3	14
11	1	4	2	4	5	1	-4	-2	4	5	4
12	2	4	5	1	3	2	4	5	1	3	15
13	5	1	2	3	4	5	1	2	3	4	15
14	2	1	5	3	4	2	1	5	3	4	15
15	3	2	5	1	4	3	2	5	0	4	14

Let ARL_0^+ and FAR_0^+ denote the in-control average run length and the false alarm rate for the upper one-sided Shewhart signed-rank control chart, respectively. For an upper one-sided chart we would take $UCL = 15$ since it is related to a false alarm rate of 0.0313 ($FAR_0^+ = 0.0313$) and an in-control average run length of 32 ($ARL_0^+ = 32$) - see Table 3.3. Although the in-control average run length of 32 is far from the desired value, which is generally taken to be 370 or 500, it is the best under present conditions. The false alarm rate (FAR_0) and the in-control average run length (ARL_0) for the symmetric two-sided Shewhart signed-rank chart can be obtained through the relationships $FAR_0 = 2FAR_0^+$ and $ARL_0 = \frac{ARL_0^+}{2}$, respectively (see Bakir (2003)). A symmetric two-sided chart is obtained by choosing $LCL = -UCL$. We take $UCL = 15$ for the two-sided Shewhart signed-rank chart,

* See SAS Program 5 in Appendix B for the calculation of the values in Table 3.2.

since it is related to a false alarm rate of 0.0626 ($FAR_0 = 2FAR_0^+ = 2\times 0.0313 = 0.0626$) and an in-control average run length of 16 ($ARL_0 = \frac{ARL_0^+}{2} = \frac{32}{2} = 16$). The two-sided signed-rank chart is shown in Figure 3.1 with $UCL = 15$, $CL = 0$ and $LCL = -15$.

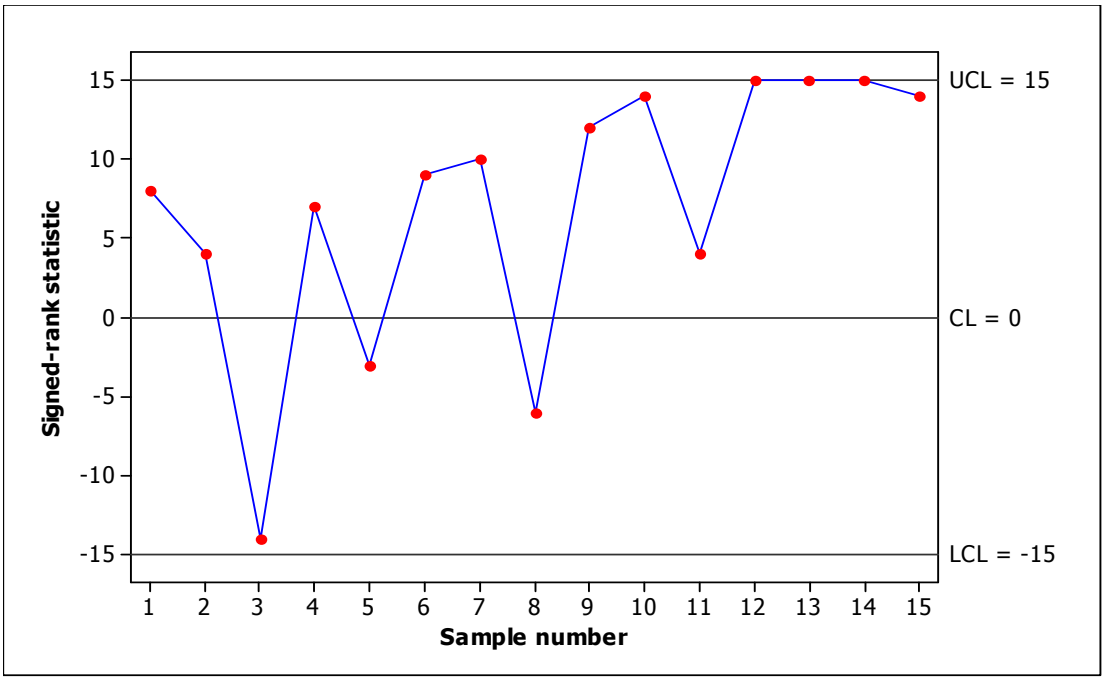

Figure 3.1. Signed-rank control chart for Montgomery (2001) piston ring data.

The chart signals at sample number 12. Therefore, a search for assignable causes is necessary. It appears most likely that the process median has shifted upwards from the target value of 74*mm*.

3.1.4. Determination of chart constants

The control limits in example 3.1 were chosen to give a certain false alarm rate or in-control ARL. Values of various control limits are given by Bakir (2003). Bakir included the following table in his article which gives the false alarm rates and the in-control average run lengths for the upper one-sided Shewhart signed-rank charts based on subgroups of sizes $n = 4,\ 5$ and 6.

Table 3.3. *FAR*'s and ARL_0's for the upper one-sided Shewhart signed-rank chart.

UCL	$n = 4$		$n = 5$		$n = 6$	
	ARL_0^+	FAR_0^+	ARL_0^+	FAR_0^+	ARL_0^+	FAR_0^+
10	16.00	0.0625	10.66	0.0938	6.40	0.1563
11	∞	0	10.66	0.0938	6.40	0.1563
12			16.00	0.0630	9.14	0.1094
13			16.00	0.0630	9.14	0.1094
14			32.00	0.0313	12.80	0.0781
15			32.00	0.0313	12.80	0.0781
16			∞	0	21.33	0.0469
17					21.33	0.0469
18					32.00	0.0312
19					32.00	0.0312
20					64.00	0.0156
21					64.00	0.0156
22					∞	0

Table 3.3 shows the false alarm rates and the in-control average run lengths for the upper one-sided Shewhart signed-rank chart as calculated using the null distribution of the Wilcoxon signed-rank statistic (see Hollander and Wolfe (1973) and Bakir (2003)).

In Table 3.3 we see that there are some duplicates in the data. We consider a specific example to shed light on the occurrence of these duplicates. Suppose $n = 5$ and $UCL = 12$. Then $FAR_0^+ = P(SR_i \geq 12 \mid \text{In-control}) = P(T_n^+ \geq 13.5)$ (using (3.2)). The last probability equals $P(T_n^+ \geq 14) = 0.0630$, because T_n^+ has zero probability at 13.5. When $n = 5$ and $UCL = 13$ we have that $FAR_0^+ = P(SR_i \geq 13 \mid \text{In-control}) = P(T_n^+ \geq 14) = 0.0630$ (by using the null distribution of the Wilcoxon signed-rank statistic). Since $FAR_0^+ = P(T_n^+ \geq 14) = 0.0630$ for two different values of the upper control limit, we have duplicates in the data. This example points out an error* in Table 1 of Bakir (2003). The probability of $P(T_n^+ \geq 13.5)$ equals $P(T_n^+ \geq 14)$ which equals 0.0630 (and not 0.0938 corresponding to $P(T_n^+ \geq 13)$ as reported by Bakir's (2003) paper). This type of correction was applied to the other entries of Bakir's (2003) Table 1 and are given in Table 3.3 of this thesis. The false alarm rates and in-control average run lengths for the two-sided Shewhart signed-rank chart were calculated using SAS (with the appropriate corrections made) and are shown in Table 3.4.

* This error is also pointed out by Chakraborti and Eryilmaz (2007).

Table 3.4. *FAR*'s and ARL_0's for the two-sided Shewhart signed-rank chart*.

UCL	*n* = 4		*n* = 5		*n* = 6		*n* = 7		*n* = 8		*n* = 9		*n* = 10	
	ARL_0	*FAR*	ARL_0	*FAR*	ARL_0	*FAR*	ARL_0	*FAR*	ARL_0	*FAR*	ARL_0	*FAR*	ARL_0	*FAR*
11	8.018	0.125	3.248	0.308	2.307	0.433	2.132	0.469	1.830	0.546	1.526	0.655	1.447	0.691
12	∞	0	5.382	0.186	3.163	0.316	2.126	0.470	1.811	0.552	1.779	0.562	1.606	0.623
13			5.392	0.185	3.234	0.309	2.643	0.378	2.146	0.466	1.730	0.578	1.613	0.620
14			8.003	0.125	4.557	0.219	2.675	0.374	2.194	0.456	2.022	0.494	1.802	0.555
15			7.896	0.127	4.623	0.216	3.360	0.298	2.597	0.385	1.991	0.502	1.788	0.559
16			15.922	0.063	6.302	0.159	3.399	0.294	2.616	0.382	2.354	0.425	2.033	0.492
17			∞	0	6.483	0.154	4.552	0.220	3.201	0.312	2.346	0.426	2.032	0.492
18					10.797	0.093	4.553	0.220	3.276	0.305	2.770	0.361	2.276	0.439
19					10.762	0.093	6.514	0.154	4.004	0.250	2.778	0.360	2.345	0.426
20					16.291	0.061	6.307	0.159	3.999	0.250	3.350	0.299	2.635	0.379
21					15.982	0.063	9.164	0.109	5.053	0.198	3.309	0.302	2.664	0.375
22					29.890	0.033	9.152	0.109	5.149	0.194	4.034	0.248	3.071	0.326
23					∞	0	12.655	0.079	6.611	0.151	4.055	0.247	3.139	0.319
24							12.618	0.079	6.632	0.151	5.055	0.198	3.641	0.275
25							20.939	0.048	9.244	0.108	5.047	0.198	3.682	0.272
26							21.115	0.047	9.299	0.108	6.032	0.166	4.355	0.230
27							31.380	0.032	12.862	0.078	6.053	0.165	4.226	0.237
28							31.118	0.032	12.898	0.078	7.768	0.129	5.225	0.191
29							64.444	0.016	18.447	0.054	7.812	0.128	5.108	0.196
30							∞	0	17.947	0.056	10.180	0.098	6.286	0.159
31									25.216	0.040	10.554	0.095	6.315	0.158
32									25.285	0.040	13.573	0.074	7.670	0.130
33									42.248	0.024	13.357	0.075	7.638	0.131
34									42.872	0.023	18.499	0.054	9.505	0.105
35									63.492	0.016	18.409	0.054	9.728	0.103
36									64.492	0.016	25.763	0.039	12.004	0.083
37									129.711	0.008	25.676	0.039	11.531	0.087
38									∞	0	37.023	0.027	15.663	0.064
39											36.507	0.027	15.514	0.064
40											50.919	0.020	20.504	0.049
41											51.913	0.019	20.542	0.049

* See SAS Program 6 in Appendix B for the calculation of the values in Table 3.4. This table is an extension of Tables 1 and 2 given in Bakir (2003).

42											87.428	0.011	27.510	0.036
43											84.898	0.012	26.771	0.037
44											127.219	0.008	36.928	0.027
45											128.950	0.008	37.308	0.027
46											251.312	0.004	50.234	0.020
47											∞	0	52.249	0.019
48													73.736	0.014
49													74.261	0.013
50													104.300	0.010
51													101.973	0.010
52													165.381	0.006
53													168.821	0.006
54													251.693	0.004
55													249.627	0.004
56													443.132	0.002
57													∞	0

3.1.5. Summary

The signed-rank test is a popular nonparametric test for the median of a symmetric continuous population. The signed-rank test is more powerful than the sign test, but while the sign test is applicable for all continuous distributions, the assumption of symmetry must be made, in addition, for the signed-rank test. Furthermore, the sign test applies to all percentiles, whereas the signed-rank test is proposed only for the median. Another drawback of the signed-rank chart is that the *FAR* values for the chart are too high (in other words the ARL_0 values are too short) unless the subgroup size is 'large'. One way to remedy this problem is to use some signaling rules to enhance the sensitivity of the charts. This will be considered next.

3.2. The Shewhart-type control chart with runs-type signaling rules

3.2.1. Introduction

In addition to defining warning limits or zones on control charts (see Section 2.2), we can extend the existing charts by incorporating various signaling rules involving runs of the plotting statistic. The signaling rules considered include the following: A process is declared to be out-of-control when (a) a single point (charting statistic) plots outside the control limit(s) (*1-of-1* rule) (b) k consecutive points (charting statistics) plot outside the control limit(s) (*k-of-k* rule) or (c) exactly k of the last w points (charting statistics) plot outside the control limit(s) (*k-of-w* rule). We can consider these signaling rules where both k and w are positive integers with $1 \le k \le w$ and $w \ge 2$. Rule (a) is the simplest and is the most frequently used in the literature. Thus, the *1-of-1* rule corresponds to the usual control chart, where a signal is given when a plotting statistic falls outside the control limit(s). Rules (a) and (b) are special cases of rule (c); rules (b) and (c) have been used in the context of supplementing the Shewhart charts with warning limits and zones. Rules (a), (b) and (c) have been studied by various authors (see for example Klein (2000) and Khoo (2004)). Klein (2000) suggested two rules namely the *2-of-2* and *2-of-3* rules. Both control charts are easily implemented and have better *ARL* performance than the *1-of-1* rule. Khoo (2004) conducted a study of the *ARL* performance of the *2-of-2*, *2-of-3*, *2-of-4*, *3-of-3* and *3-of-4* charts and concluded that the *3-of-4* chart is the most sensitive scheme for detecting small process shifts.

Chakraborti and Eryilmaz (2007) considered simple alternatives to the Bakir (2004)'s class of nonparametric charts, using the signed-rank statistic but incorporating runs rules of the type discussed above to define new signaling rules. If we set k equal to 2 in rule (b) above, we obtain the simplest of the *k-of-k* type rules which are called the *2-of-2* DR and the *2-of-2* KL charts. The *2-of-2* KL chart signals, for example, when the two most recent signed-rank statistics both fall either on or above or on or below the control limits. The *2-of-2* DR chart is almost similar, but here a signal is indicated when both of the signed-rank statistics fall either both on or above or both on or below or one on or above (below) and the next one on or below (above) the control limits. It is shown that the new charts are nonparametric, have much smaller *FAR* (and thus larger ARL_0) than the *1-of-1* signed-rank chart of Bakir. Moreover, the new charts have better out-of-control performance than the *1-of-1* signed-rank chart for heavy-tailed and skewed distributions such as the Cauchy. We illustrate these procedures using the Montgomery (2001) piston ring data.

3.2.2. Example

Example 3.2

A two-sided Shewhart signed-rank chart with signaling rules for the Montgomery (2001) piston ring data

We illustrate the signed-rank chart with signaling rules using the Montgomery (2001) piston ring data. Recall that the dataset contains 15 samples (each of size 5). The signed-rank statistics were calculated and given in Table 3.2 and graphically represented in Figure 3.1. The symmetric two-sided control limits for the *1-of-1* and *2-of-2* signed-rank charts are given by Chakraborti and Eryilmaz (2007) for $n = 4, 5, 6$ and 10. The table for samples of size 5 is given here for reference.

Table 3.5. False alarm rates and in-control *ARL* values for the two-sided *1-of-1* and *2-of-2* signed-rank charts under DR and KL schemes, $n = 5$ [*].

	1-of-1		*2-of-2* DR		*2-of-2* KL	
UCL	ARL_0	FAR_0	$ARL_{DR,0}$	$FAR_{DR,0}$	$ARL_{KL,0}$	$FAR_{KL,0}$
11	5.33	0.1876	33.74	0.0352	62.15	0.0176
13	8.00	0.1250	72.00	0.0156	136.00	0.0078
15	15.97	0.0626	271.15	0.0039	526.34	0.0019

For $n = 5$, the control limits for the *1-of-1* (Bakir's chart), the *2-of-2* DR and the *2-of-2* KL charts, based on the signed-rank statistic, are set at ± 15. These yield *FAR* values 0.0626, 0.0039, and 0.0019, respectively. If the control limits were taken to be ± 13, the *FAR* would have been much higher: 0.1250, 0.0156, and 0.0078, respectively. Although the control limits are the same, namely ± 15, the signaling rules are quite different operationally and the performance of the resulting charts turn out to be quite different. The *1-of-1* chart signals when the first signed-rank statistic falls on or outside of either of the two control limits; the *2-of-2* KL chart signals when, for the first time, two consecutive signed-rank statistics fall either on or above or on or below the two control limits, while the *2-of-2* DR chart signals when for the first time two consecutive signed-rank statistics fall on or outside the control limits, either both on or above, or both on or below, or one on or above the next on or below, or one on or below and the next on or above. On the performance side, note that the *1-of-1* SR chart has a *FAR* of 0.0626 and an ARL_0 of approximately 16. Thus many more false alarms will be signaled by this chart leading to a possible loss of time and resources. Compared to that, the *2-of-2* KL chart has a *FAR* of 0.0019 and an ARL_0 of 526.34, whereas the *2-of-2* DR chart has a *FAR* of 0.0039 and an ARL_0 of 271.15. Thus both of these run-rule-enhanced charts provide reasonable and practical false alarm rates and can be used in practice, depending on the type of shift one expects.

From Figure 3.1 we see that the DR and KL *2-of-2* signed-rank charts both signal at sample 13, indicating a most likely upward shift in the process median. The *1-of-1* signed-rank chart, on the other hand, signals earlier, at sample 12, but note the much higher *FAR* of 0.0626 (and correspondingly a much lower and less desirable ARL_0, 15.97) associated with this chart. It is interesting to note that, as shown in Montgomery (2001), for these data the

[*] Table 3.5 appears in Chakraborti and Eryilmaz (2007), Table 11.

Shewhart $\overline{X}$ chart indicates a shift in the mean at sample 11 for these data. However, the key difference is that an application of the Shewhart chart can raise several questions such as the form of the underlying distribution (small $n = 5$), and more importantly about the in-control (stable) performance of the chart in terms of the *FAR* (or the ARL_0), since it is known that the in-control performance of the Shewhart $\overline{X}$ chart is not robust in typical quality control applications. Compared to this, the proposed nonparametric charts provide a more generally applicable alternative monitoring scheme with a known (stable/robust) in-control performance and a better or equal out-of-control performance than the *1-of-1* signed-rank chart.

3.2.3. Summary

In this section we examined signed-rank control charts with runs-type signaling rules. Human, Chakraborti and Smit (2008) recently studied Shewhart-type sign charts with runs-type signaling rules. These charts are similar in spirit to the Shewhart-type signed-rank charts with runs-type signaling rules (see Section 3.2). In the paper by Human et al. they derived expressions for the run length distributions using Markov chain theory. The in-control and out-of-control performance of these charts were studied and compared to those of the existing signed-ranked charts under the normal, double exponential and Cauchy distributions, using the *ARL*, *SDRL*, *FAR* and some percentiles of the run length. These runs rules enhanced sign charts have the advantage that one does not have to assume symmetry of the underlying distribution and they can be applied in situations where the data are dichotomous.

3.3. The tabular CUSUM control chart

3.3.1. Introduction

Bakir and Reynolds (1979) investigated the CUSUM chart using the Wilcoxon signed-rank statistic. They used methods that are analogous to the methods used on the CUSUM sign chart (see Section 2.3), that is, a Markov chain approach is used to find the moments and other characteristics of the run length distribution for the CUSUM signed-rank chart.

3.3.2. One-sided control charts

3.3.2.1. Upper one-sided control charts

Fu, Spiring and Xie (2002) and Fu and Lou (2003) presented three results that must be satisfied before implementing the finite-state Markov chain approach. Let S_t^+ be a finite-state homogenous Markov chain on the state space Ω^+ with a transition probability matrix (TPM) such that (i) $\Omega^+ = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$ where $0 = \varsigma_0 < \varsigma_1 < ... < \varsigma_{r+s-1} = h$ and ς_{r+s-1} is an absorbent state; (ii) the TPM is given by $TPM = [p_{ij}]$ for $i = 0,1,...,r+s-1$ and $j = 0,1,...,r+s-1$ where r denotes the number of non-absorbent* states and s the number of absorbent† states, respectively, and (iii) the starting value should equal zero with probability one, that is, $P(S_0^+ = 0) = 1$ (this is to ensure that the process starts in-control). Assume that the Markov chain S_t^+ satisfies conditions (i), (ii) and (iii), then the formulas given in (2.41) to (2.45) hold.

The time that the procedure signals is the first time such that the finite-state Markov chain S_t^+ enters the state ς_{r+s-1} where the state space is given by $\Omega^+ = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$, $S_0^+ = 0$ and

$$S_t^+ = \min\{h, \max\{0, S_{t-1}^+ + SR_t - k\}\} \tag{3.3}$$

* The transient (non-absorbent) states are the states for which eventual return is uncertain.
† If a state is entered once and is never left, the state is said to be absorbent.

3.4. The EWMA control chart

3.4.1. Introduction

In this section, the approach taken by Lucas and Saccucci (1990) is extended to the use of the signed-rank statistic resulting in an EWMA signed-rank chart that accumulates the statistics $SR_1, SR_2, SR_3, \ldots$. Section 3.4 is analogous to Section 2.4 where the approach taken by Lucas and Saccucci (1990) was extended to the use of the sign statistic resulting in an EWMA sign chart. Therefore, the reader is frequently referred back to Section 2.4 throughout this section.

3.4.2. The proposed EWMA signed-rank chart

A nonparametric EWMA-type of control chart based on the signed-rank statistic (recall that $SR_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0) R_{ij}^{+}$) can be obtained by replacing X_i in expression (2.53) of Section 2.4 with SR_i. The EWMA signed-rank chart accumulates the statistics $SR_1, SR_2, SR_3, \ldots$ with the plotting statistics defined as

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

where $0 < \lambda \leq 1$ is a constant called the weighting constant. The starting value Z_0 could be taken to equal zero, i.e. $Z_0 = 0$.

The EWMA signed-rank chart is constructed by plotting Z_i against the sample number i (or time). If the plotting statistic Z_i falls between the two control limits, that is, $LCL < Z_i < UCL$, the process is considered to be in-control. If the plotting 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 out-of-control.

The exact control limits and the center line of the EWMA signed-rank control chart can be obtained by replacing σ and θ_0 by σ_{SR_i} and 0, respectively, in expression (2.55) of Section 2.4 to obtain

$$\begin{aligned} UCL &= L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \\ CL &= 0 \\ LCL &= -L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \end{aligned} \quad . \tag{3.11}$$

Similarly, the *steady-state* control limits can be obtained by replacing σ and θ_0 by σ_{SR_i} and 0, respectively, in expression (2.56) to obtain

$$\begin{aligned} UCL &= L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \\ LCL &= -L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \end{aligned} \tag{3.12}$$

where σ_{SR_i} denotes the in-control standard deviation of the signed-rank statistic SR_i if there are no ties within a subgroup.

The in-control standard deviation of SR_i is given by $\sigma_{SR_i} = \sqrt{\text{var}(SR_i)} = \sqrt{\text{var}\left(2T^+ - \frac{n(n+1)}{2}\right)} = \sqrt{\frac{n(n+1)(2n+1)}{6}}$. This is obtained by using the relationship between SR_i and T^+ (recall that $SR_i = 2T^+ - \frac{n(n+1)}{2}$ if there are no ties within a subgroup) and the fact that $\text{var}(T^+) = \frac{n(n+1)(2n+1)}{24}$ (see Gibbons and Chakraborti (2003) page 198).

3.4.3. Markov-chain approach

Lucas and Saccucci (1990) evaluated the properties of the *continuous state* Markov chain by *discretizing* the infinite state TPM. This procedure entails dividing the interval between the UCL and the LCL into N subintervals of width 2δ. Then the plotting statistic, Z_i, is said to be in the non-absorbing state j at time i if $S_j - \delta < Z_i \leq S_j + \delta$ where S_j denotes the midpoint of the j^{th} interval. Z_i is said to be in the absorbing state if Z_i falls on or outside one of the control

limits, that is, $Z_i \le LCL$ or $Z_i \ge UCL$. Let p_{ij} denote the probability of moving from state i to state j in one step, i.e. $p_{ij} = P(Moving\ to\ state\ j \mid in\ state\ i)$. To approximate this probability we assume that the plotting statistic is equal to S_i whenever it is in state i. For all j non-absorbing we obtain $p_{ij} = P(S_j - \delta < Z_k \le S_j + \delta \mid Z_{k-1} = S_i)$. By using the definition of the plotting statistic given in expression (3.10) we obtain

$$\begin{aligned} p_{ij} &= P(S_j - \delta < \lambda SR_k + (1-\lambda)S_i \le S_j + \delta) \\ &= P\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} < SR_k \le \frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda}\right) \end{aligned}$$

recall that $SR_k = 2T_k^+ - \frac{n(n+1)}{2}$

$$\begin{aligned} p_{ij} &= P\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} < 2T_k^+ - \frac{n(n+1)}{2} \le \frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(\frac{\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}\right)}{2} < T_k^+ \le \frac{\left(\frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}\right)}{2}\right). \end{aligned} \tag{3.13}$$

For all j absorbing we obtain

$$\begin{aligned} p_{ij} &= P(Z_k \le LCL \mid Z_{k-1} = S_i) + P(Z_k \ge UCL \mid Z_{k-1} = S_i) \\ &= P(\lambda SR_k + (1-\lambda)S_i \le LCL) + P(\lambda SR_k + (1-\lambda)S_i \ge UCL) \\ &= P\left(SR_k \le \frac{LCL - (1-\lambda)S_i}{\lambda}\right) + P\left(SR_k \ge \frac{UCL - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(2T_k^+ - \frac{n(n+1)}{2} \le \frac{LCL - (1-\lambda)S_i}{\lambda}\right) + P\left(2T_k^+ - \frac{n(n+1)}{2} \ge \frac{UCL - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(T_k^+ \le \frac{\frac{LCL - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}}{2}\right) + P\left(T_k^+ \ge \frac{\frac{UCL - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}}{2}\right). \end{aligned} \tag{3.14}$$

Since the values LCL, UCL, δ, λ, n, S_i and S_j are known constants the Wilcoxon signed-rank probabilities in expressions (3.13) and (3.14) can easily be calculated. The probabilities for the Wilcoxon signed-rank statistics are given in Table H of Lehmann (1975) for

samples sizes up to 20 and they are tabulated (more recently) in Table H of Gibbons and Chakraborti (2003) for sample sizes up to 15.

Once the one-step transition probabilities are calculated, the TPM can be constructed and is given by $TPM = [p_{ij}] = \begin{pmatrix} Q & | & \underline{p} \\ - & - & - \\ \underline{0}' & | & 1 \end{pmatrix}$ (written in partitioned form) where the essential transition probability sub-matrix Q is the matrix that contains all the transition probabilities of going from a non-absorbing state to a non-absorbing state, $Q:(NA \rightarrow NA)$, $\underline{p}$ contains all the transition probabilities of going from each non-absorbing state to the absorbing states, $\underline{p}:(NA \rightarrow A)$, $\underline{0}' = (0 \quad 0 \quad 0 \quad \cdots \quad 0)$ contains all the transition probabilities of going from each absorbing state to the non-absorbing states. $\underline{0}'$ is a row vector with all its elements equal to zero, because it is impossible to go from an absorbing state to a non-absorbing state, because once an absorbing state is entered, it is never left, $\underline{0}' :(A \rightarrow NA)$, and 1 represents the scalar value one. The probability of going of going from an absorbing state to an absorbing state is equal to one, because once an absorbing state is entered, it is never left, $1:(A \rightarrow A)$. The one-step TPM is used to calculate the expected value (*ARL*), the second raw moment, the variance, the standard deviation and the probability mass function (pmf) of the run-length variable N which are given in equations (2.41) to (2.45).

Example 3.10

The EWMA signed-rank chart where the sample size is even ($n = 6$)

The EWMA signed-rank chart is investigated for a smoothing constant of 0.1 ($\lambda = 0.1$) and a multiplier of 3 ($L = 3$). The *steady-state* control limits are given by

$$UCL = L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

$$LCL = -L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

where $L = 3$, $\lambda = 0.1$, and $\sigma_{SR_i} = 9.539$, since $\sigma_{SR_i} = \sqrt{\frac{n(n+1)(2n+1)}{6}} = \sqrt{\frac{6(6+1)(12+1)}{6}} = 9.539$. Clearly, we only have to calculate the UCL since $LCL = -UCL$. We obtain $UCL = 3 \times 9.539\sqrt{\left(\frac{0.1}{2-0.1}\right)} = 6.565$. Therefore, $LCL = -6.565$.

This Markov-chain procedure entails dividing the interval between the UCL and the LCL into N subintervals of width 2δ. For this example N is taken to equal 4. Figure 3.13 illustrates the partitioning of the interval between the UCL and the LCL into subintervals.

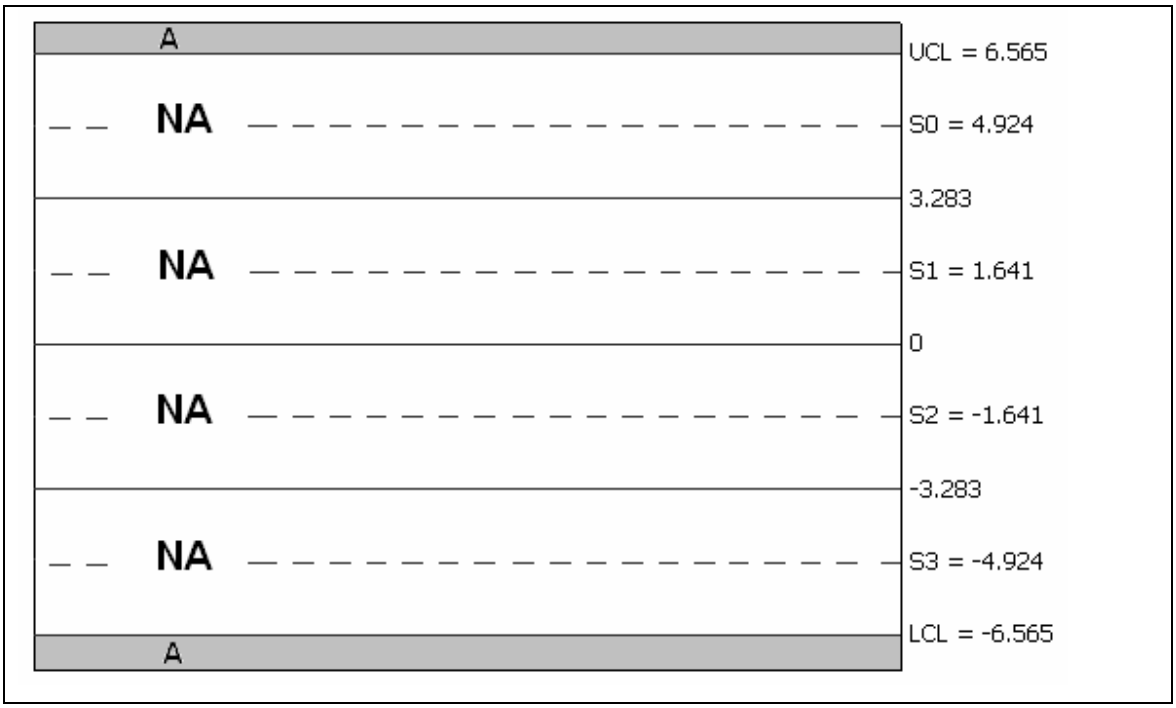

Figure 3.13. Partitioning of the interval between the UCL and the LCL into 4 subintervals.

From Figure 3.13 we see that there are 4 non-absorbing states, i.e. $r = 4$. The TPM is given by

$$TPM = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \begin{pmatrix} Q_{4\times4} & | & \underline{p}_{4\times1} \\ - & - & - \\ \underline{0}'_{1\times4} & | & 1_{1\times1} \end{pmatrix}.$$

Table 3.31. Calculation of the one-step probabilities in the first row of the TPM.

$p_{00} = P(\textit{Moving to state } 0 \mid \textit{in state } 0)$

$= P(S_0 - \delta < Z_k \le S_0 + \delta \mid Z_{k-1} = S_0)$ from (3.13)

$$= P\left(\frac{\left(\frac{(S_0 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_0 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$

with $\delta = 1.641$, $\lambda = 0.1$, $L = 3$ and $S_0 = 4.924$

$= P(4.755 < T_k^+ \le 21.169)$

$= P(T_k^+ \le 21) - P(T_k^+ \le 4)$

$= \frac{57}{64}$ from Gibbons and Chakraborti (2003)

$p_{01} = P(\textit{Moving to state } 1 \mid \textit{in state } 0)$

$= P(S_1 - \delta < Z_k \le S_1 + \delta \mid Z_{k-1} = S_0)$ from (3.13)

$$= P\left(\frac{\left(\frac{(S_1 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_1 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$

$= P(-11.658 < T_k^+ \le 4.755)$

$= P(T_k^+ \le 4)$

$= \frac{7}{64}$

$$p_{02} = P(\textit{Moving to state } 2 \mid \textit{in state } 0)$$
$$= P(S_2 - \delta < Z_k \le S_2 + \delta \mid Z_{k-1} = S_0) \text{ from (3.13)}$$
$$= P\left(\frac{\left(\frac{(S_2 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_2 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$
$$= P(-28.072 < T_k^+ \le -11.658)$$
$$= 0$$

$$p_{03} = P(\textit{Moving to state } 3 \mid \textit{in state } 0)$$
$$= P(S_3 - \delta < Z_k \le S_3 + \delta \mid Z_{k-1} = S_0) \text{ from (3.13)}$$
$$= P\left(\frac{\left(\frac{(S_3 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_3 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$
$$= P(-44.486 < T_k^+ \le -28.072)$$
$$= 0$$

$$p_{04} = P(\textit{Moving to state } 4 \mid \textit{in state } 0)$$
$$= P(Z_k \le LCL \mid Z_{k-1} = S_0) + P(Z_k \ge UCL \mid Z_{k-1} = S_0) \text{ from (3.14)}$$
$$= P\left(T_k^+ \le \frac{\frac{LCL - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2}}{2} \right) + P\left(T_k^+ \ge \frac{\frac{UCL - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2}}{2} \right)$$
$$= P(T_k^+ \le -44.486) + P(T_k^+ \ge 21.169)$$
$$= 0$$

The one-step probabilities in the remaining rows can be calculated similarly. Therefore, the TPM is given by $TPM = \begin{pmatrix} 57/64 & 7/64 & 0 & 0 & 0 \\ 2/64 & 57/64 & 5/64 & 0 & 0 \\ 0 & 5/64 & 57/64 & 2/64 & 0 \\ 0 & 0 & 7/64 & 57/64 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_{4\times4} & | & \underline{p}_{4\times1} \\ - & - & - \\ \underline{0}'_{1\times4} & | & 1_{1\times1} \end{pmatrix}$.

Other values of the multiplier (L) and the smoothing constant (λ) were also considered and the results are given in Tables 3.32 and 3.33.

Table 3.32. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for the EWMA signed-rank chart when $n = 6$ and $N = 5$, i.e. there are 5 subintervals between the lower and upper control limit†.

	$L = 1$	$L = 2$	$L = 3$
$\lambda = 0.05$	10.45 12.32 (1, 2, 6, 15, 35)	56.69 72.45 (1, 5, 29, 82, 204)	**
$\lambda = 0.1$	7.32 8.38 (1, 1, 4, 10, 24)	33.83 40.28 (1, 4, 20, 48, 115)	330.67 369.33 (2, 63, 213, 471, 1070)
$\lambda = 0.2$	4.95 4.90 (1, 1, 3, 7, 15)	35.21 39.63 (1, 6, 22, 50, 115)	361.92 384.29 (3, 87, 243, 510, 1130)

** The inverse of the matrix $(I - Q)$ does not exist and as a result the ARL (given by $E(N) = \underline{\xi}(I - Q)^{-1}\underline{1}$) can not be calculated for this combination of (λ, L).

In example 3.10 we considered a sample size that may be considered "small". The results are given for a larger sample size ($n = 10$) for various values of λ and L in Table 3.33.

Table 3.33. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values‡ for the EWMA signed-rank chart when $n = 10$ and $N = 5$, i.e. there are 5 subintervals between the lower and upper control limit§.

	$L = 1$	$L = 2$	$L = 3$
$\lambda = 0.05$	11.17 13.49 (1, 2, 6, 16, 39)	67.94 83.82 (1, 7, 38, 98, 238)	1448.44 1573.37 (10, 316, 956, 2052, 4595)
$\lambda = 0.1$	6.85 7.74 (1, 1, 4, 9, 23)	48.87 57.73 (1, 6, 29, 70, 165)	352.72 384.51 (3, 76, 232, 500, 1122)
$\lambda = 0.2$	5.05 5.07 (1, 1, 3, 7, 15)	33.96 38.48 (1, 6, 21, 48, 111)	336.34 357.54 (3, 80, 226, 474, 1051)

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.
† See SAS Program 8 in Appendix B for the calculation of the values in Table 3.32.
‡ The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.
§ See SAS Program 8 in Appendix B for the calculation of the values in Table 3.33.

These tables can be extended by changing the sample size (n), the number of subintervals between the lower and upper control limit (N), the multiplier (L) and the smoothing constant (λ) in SAS Program 8 for the EWMA signed-rank chart given in Appendix B.

From Tables 3.32 and 3.33 we see that the ARL_0, $SDRL$ and percentiles increase as the value of the multiplier (L) increases. From Table 3.33 we find an in-control average run length of 336.34 for $n = 10$ when the multiplier is taken to equal 3 ($L = 3$) and the smoothing constant 0.2 ($\lambda = 0.2$). The chart performance is good, since the attained in-control average run length of 336.34 is in the region of the desired in-control average run length which is generally taken to be 370 or 500.

3.4.4. Summary

The EWMA control chart is one of several charting methods aimed at correcting a deficiency of the Shewhart chart - insensitivity to small shifts. Lucas and Saccucci (1990) have investigated some properties of the EWMA chart under the assumption of independent normally distributed observations, whereas in this section we have described and evaluated the nonparametric EWMA signed-rank chart. The main advantage of the nonparametric EWMA chart is that there is no need to assume a particular parametric distribution for the underlying process (see Section 1.4 for other advantages of the nonparametric EWMA chart).

where h is the decision interval and k is the reference value (see Section 2.3.1 for a detailed discussion on how the values of k and h are chosen). Equation (3.3) is obtained by replacing SN_t with SR_t in (2.46).

The distribution of SR_t can easily be obtained from the distribution of the Wilcoxon signed-rank statistic T^+ (recall that $SR_i = 2T_i^+ - \frac{n(n+1)}{2} \quad \forall i$). The probabilities for the Wilcoxon signed-rank statistics are given in Table H of Lehmann (1975) for samples sizes up to 20 and they are tabulated (more recently) in Table H of Gibbons and Chakraborti (2003) for sample sizes up to 15.

Example 3.3

An upper one-sided CUSUM signed-rank chart where the sample size is even (n=4)

The statistical properties of an upper one-sided CUSUM signed-rank chart with a decision interval of 6 $(h = 6)$, a reference value of 2 $(k = 2)$ and a sample size of 4 $(n = 4)$ is examined. We start by examining the pmf of the well-known Wilcoxon signed-rank statistic T^+, since the plotting statistic SR_i is linearly related to T^+.

Table 3.6. Enumeration for the distribution of T^+ for a sample size of 4.

Value of T^+	Ranks associated with positive differences	Number of sample points $u(t)$	$P(T^+ = t)$	$P(T^+ \leq t)$
0		1	1/16	1/16
1	1	1	1/16	2/16
2	2	1	1/16	3/16
3	{1,2}; {3}	2	2/16	5/16
4	{1,3}; {4}	2	2/16	7/16
5	{1,4}; {2,3}	2	2/16	9/16
6	{1,2,3}; {2,4}	2	2/16	11/16
7	{1,2,4}; {3,4}	2	2/16	13/16
8	{1,3,4}	1	1/16	14/16
9	{2,3,4}	1	1/16	15/16
10	{1,2,3,4}	1	1/16	16/16

From Table 3.6 if follows that the pmf of T^+ when the sample size is 4 is

$$f_{T^+}(t) = P(T^+ = t) = \begin{cases} 1/16 & t = 0,1,2,8,9,10 \\ 2/16 & t = 3,4,5,6,7 \\ 0 & \text{otherwise} \end{cases}$$

The values of SR_t are either the even or the odd integers between (and including) $-\frac{n(n+1)}{2}$ and $\frac{n(n+1)}{2}$, depending on whether $\frac{n(n+1)}{2}$ is even or odd. In example 3.3 $\frac{n(n+1)}{2} = \frac{4(4+1)}{2} = 10$ which is even and as a result the possible values for SR_t are even integers between -10 and 10 inclusive. Thus, we have that $-10 \leq SR_t \leq 10$. In both cases (whether $\frac{n(n+1)}{2}$ is even or odd) the sum $\sum(SR_i - k)$ will be an integer since both SR_t and k are integers. For this example, the reference value is taken to be equal to two, because this leads to the sum $\sum(SR_i - k)$ being equal to even values which reduces the size of the state space for the Markov chain. For $h = 6$ we have that $\Omega^+ = \{\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3\} = \{0,2,4,6\}$ with

$0 = \varsigma_0 < \varsigma_1 < \varsigma_2 < \varsigma_3 = h$. The state space is calculated using equation (3.3) and the calculations are shown in Table 3.7.

Table 3.7. Calculation of the state space when $h = 6$, $k = 2$ and $n = 4$.

SR_t	$S_{t-1}^+ + SR_t - k$	$\max\{0, S_{t-1}^+ + SR_t - k\}$	$S_t^+ = \min\{h, \max\{0, S_{t-1}^+ + SR_t - k\}\}$
-10	-12*	0	0
-8	-10	0	0
-6	-8	0	0
-4	-6	0	0
-2	-4	0	0
0	-2	0	0
2	0	0	0
4	2	2	2
6	4	4	4
8	6	6	6
10	8	8	6

Table 3.8. Classification of the states.

State number	Description of the state	Absorbent (A)/ Non-absorbent (NA)
0	$S_t^+ = 0$	NA
1	$S_t^+ = 2$	NA
2	$S_t^+ = 4$	NA
3	$S_t^+ = 6$	A

From Table 3.8 we see that there are three non-absorbent states, i.e. $r = 3$, and one absorbent state, i.e. $s = 1$. Therefore, the corresponding TPM will be a $(r+s)\times(r+s) = 4\times 4$ matrix. It can be shown (see Table 3.9) that the TPM is given by

$$TPM_{4\times 4} = \begin{pmatrix} p_{00} & p_{02} & p_{04} & p_{06} \\ p_{20} & p_{22} & p_{24} & p_{26} \\ p_{40} & p_{42} & p_{44} & p_{46} \\ p_{60} & p_{62} & p_{64} & p_{66} \end{pmatrix} = \left(\begin{array}{cccc|c} 11/16 & 2/16 & 1/16 & & 2/16 \\ 9/16 & 2/16 & 2/16 & & 3/16 \\ 7/16 & 2/16 & 2/16 & & 5/16 \\ - & - & - & - & - \\ 0 & 0 & 0 & & 1 \end{array}\right) = \begin{pmatrix} Q_{3\times 3} & \vert & \underline{p}_{3\times 1} \\ - & - & - \\ \underline{0}'_{1\times 3} & \vert & 1_{1\times 1} \end{pmatrix}$$

* Note: Since only the state space needs to be described, S_{t-1}^+ can be any value from Ω^+ and we therefore take, without loss of generality, $S_{t-1}^+ = 0$. Any other possible value for S_{t-1}^+ would lead to the same Ω^+.

where the essential transition probability sub-matrix $Q_{3\times3}:(NA\rightarrow NA)$ is an $r\times r=3\times3$ matrix, $\underline{p}_{3\times1}:(NA\rightarrow A)$ is an $(r+s-1)\times1=3\times1$ column vector, $\underline{0}'_{1\times3}:(A\rightarrow NA)$ is a $1\times(r+s-1)=1\times3$ row vector and $1_{1\times1}:(A\rightarrow A)$ represents the scalar value one.

The one-step transition probabilities are calculated by substituting SR_t in expression (3.3) by $2T^{+}-\frac{n(n+1)}{2}$ and substituting in values for h, k, S_t^{+} and S_{t-1}^{+}. The calculation of the one-step transition probabilities are given for illustration in Table 3.9.

The probabilities in the last column of the TPM can be calculated using the fact that $\sum_{j\in\Omega}p_{ij}=1\quad\forall i$ (see equation (2.18)). Therefore,

$p_{06}=1-(p_{00}+p_{02}+p_{04})=1-(11/16+2/16+1/16)=2/16$;

$p_{26}=1-(p_{20}+p_{22}+p_{24})=1-(9/16+2/16+2/16)=3/16$;

$p_{46}=1-(p_{40}+p_{42}+p_{44})=1-(7/16+2/16+2/16)=5/16$;

$p_{66}=1-(p_{60}+p_{62}+p_{64})=1-(0+0+0)=1$.

Table 3.9. The calculation of the transition probabilities when $h=6$, $k=2$ and $n=4$.

p_{00} $=P(S_t=0 \mid S_{t-1}=0)$ $=P(\min\{6,\max\{0,0+SR_t-2\}\}=0)$ $=P(\max\{0,SR_t-2\}=0)$ $=P(SR_t-2\le 0)$ $=P(SR_t\le 2)$ $=P(2T^+-10\le 2)$ $=P(T^+\le 6)$ $=\frac{11}{16}$	p_{02} $=P(S_t=2 \mid S_{t-1}=0)$ $=P(\min\{6,\max\{0,0+SR_t-2\}\}=2)$ $=P(\max\{0,SR_t-2\}=2)$ $=P(SR_t-2=2)$ $=P(SR_t=4)$ $=P(2T^+-10=4)$ $=P(T^+=7)$ $=\frac{2}{16}$	p_{04} $=P(S_t=4 \mid S_{t-1}=0)$ $=P(\min\{6,\max\{0,0+SR_t-2\}\}=4)$ $=P(\max\{0,SR_t-2\}=4)$ $=P(SR_t-2=4)$ $=P(SR_t=6)$ $=P(2T^+-10=6)$ $=P(T^+=8)$ $=\frac{1}{16}$
p_{20} $=P(S_t=0 \mid S_{t-1}=2)$ $=P(\min\{6,\max\{0,2+SR_t-2\}\}=0)$ $=P(\max\{0,SR_t\}=0)$ $=P(SR_t\le 0)$ $=P(2T^+-10\le 0)$ $=P(T^+\le 5)$ $=\frac{9}{16}$	p_{22} $=P(S_t=2 \mid S_{t-1}=2)$ $=P(\min\{6,\max\{0,2+SR_t-2\}\}=2)$ $=P(\max\{0,SR_t\}=2)$ $=P(SR_t=2)$ $=P(2T^+-10=2)$ $=P(T^+=6)$ $=\frac{2}{16}$	p_{24} $=P(S_t=4 \mid S_{t-1}=2)$ $=P(\min\{6,\max\{0,2+SR_t-2\}\}=4)$ $=P(\max\{0,SR_t\}=4)$ $=P(SR_t=4)$ $=P(2T^+-10=4)$ $=P(T^+=7)$ $=\frac{2}{16}$
p_{40} $=P(S_t=0 \mid S_{t-1}=4)$ $=P(\min\{6,\max\{0,4+SR_t-2\}\}=0)$ $=P(\max\{0,SR_t+2\}=0)$ $=P(SR_t+2\le 0)$ $=P(SR_t\le -2)$ $=P(2T^+-10\le -2)$ $=P(T^+\le 4)$ $=\frac{7}{16}$	p_{42} $=P(S_t=2 \mid S_{t-1}=4)$ $=P(\min\{6,\max\{0,4+SR_t-2\}\}=2)$ $=P(\max\{0,SR_t+2\}=2)$ $=P(SR_t+2=2)$ $=P(SR_t=0)$ $=P(2T^+-10=0)$ $=P(T^+=5)$ $=\frac{2}{16}$	p_{44} $=P(S_t=4 \mid S_{t-1}=4)$ $=P(\min\{6,\max\{0,4+SR_t-2\}\}=4)$ $=P(\max\{0,SR_t+2\}=4)$ $=P(SR_t+2=4)$ $=P(SR_t=2)$ $=P(2T^+-10=2)$ $=P(T^+=6)$ $=\frac{2}{16}$
p_{60} $=P(S_t=0 \mid S_{t-1}=6)$ $=0$ *	p_{62} $=P(S_t=2 \mid S_{t-1}=6)$ $=0$	p_{64} $=P(S_t=4 \mid S_{t-1}=6)$ $=0$

Using the TPM the *ARL* can be calculated using $ARL=\underline{\xi}(I-Q)^{-1}\underline{1}$. A well-known concern is that important information about the performance of a control chart can be missed when only examining the *ARL* (this is especially true when the process distribution is skewed). Various authors, see for example, Radson and Boyd (2005) and Chakraborti (2007), have

* The probability equals zero, because it is impossible to go from an absorbent state to a non-absorbent state.

suggested that one should examine a number of percentiles, including the median, to get the complete information about the performance of a control chart. Therefore, we now also consider percentiles. The $100\rho^{th}$ percentile is defined as the smallest integer l such that the cdf is at least $(100\times\rho)\%$. Thus, the $100\rho^{th}$ percentile l is found from $P(N \le l) \ge \rho$. The median (50^{th} percentile) will be considered, since it is a more representative performance measure than the *ARL*. The first and third quartiles (25^{th} and 75^{th} percentiles) will also be considered, since it contains the middle half of the distribution. The 'tails' of the distribution should also be examined and therefore the 5^{th} and 95^{th} percentiles are calculated. The calculation of these percentiles is shown below for illustration purposes.

Table 3.10. Calculation of the percentiles when $h = 6$, $k = 2$ and $n = 4$ [*].

N	$P(N \le l)$	The 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentiles
1	0.125	$\rho_{0.05} = 1$ (smallest integer such that the cdf is at least 0.05)
2	0.254	$\rho_{0.25} = 2$ (smallest integer such that the cdf is at least 0.25)
3	0.366	
4	0.462	
5	0.544	$\rho_{0.5} = 5$ (smallest integer such that the cdf is at least 0.5)
6	0.613	
7	0.671	
8	0.721	
9	0.763	$\rho_{0.75} = 9$ (smallest integer such that the cdf is at least 0.75)
10	0.799	
11	0.829	
12	0.855	
13	0.877	
14	0.896	
15	0.912	
16	0.925	
17	0.936	
18	0.946	
19	0.954	$\rho_{0.95} = 19$ (smallest integer such that the cdf is at least 0.95)
20[†]	0.961	

[*] See SAS Program 7 in Appendix B for the calculation of the values in Table 3.10.
[†] The value of the run length variable is only shown up to $N = 20$ for illustration purposes.

The formulas of the moments and some characteristics of the run length distribution have been studied by Fu, Spiring and Xie (2002) and Fu and Lou (2003) – see equations (2.41) to (2.45). By substituting $\underline{\xi}_{1\times3} = (1 \quad 0 \quad 0)$, $Q_{3\times3} = \begin{pmatrix} 11/16 & 2/16 & 1/16 \\ 9/16 & 2/16 & 2/16 \\ 7/16 & 2/16 & 2/16 \end{pmatrix}$ and $\underline{1}_{3\times1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ into these equations, we obtain the following:

$$ARL = E(N) = \underline{\xi}(I-Q)^{-1}\underline{1} = 6.81$$

$$E(N^2) = \underline{\xi}(I+Q)(I-Q)^{-2}\underline{1} = 83.64$$

$$SDRL = \sqrt{Var(N)} = \sqrt{E(N^2) - (E(N))^2} = 6.11$$

5^{th} percentile $= \rho_5 = 1$

25^{th} percentile $= \rho_{25} = 2$

Median $= 50^{th}$ percentile $= \rho_{50} = 5$

75^{th} percentile $= \rho_{75} = 9$

95^{th} percentile $= \rho_{95} = 19$

Other values of *h*, *k* and *n* were also considered and the results are given in Table 3.11.

Table 3.11. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for the upper one-sided CUSUM signed-rank chart when $n = 4$ †.

k	h				
	2	**4**	**6**	**8**	**10**
0	2.29 1.71 (1, 1, 2, 3, 6)	3.05 2.44 (1, 1, 2, 4, 8)	4.27 3.50 (1, 2, 3, 6, 11)	5.49 4.55 (1, 2, 4, 7, 15)	7.24 5.98 (1, 3, 5, 10, 19)
2	3.20 2.65 (1, 1, 2, 4, 8)	4.92 4.31 (1, 2, 4, 7, 14)	6.81 6.11 (1, 2, 5, 9, 19)	10.17 9.21 (1, 4, 7, 14, 29)	
4	5.33 4.81 (1, 2, 4, 7, 15)	7.74 7.19 (1, 3, 6, 11, 22)	13.28 12.58 (1, 4, 9, 18, 38)		
6	8.00 7.48 (1, 3, 6, 11, 23)	15.06 14.49 (1, 5, 11, 21, 44)			
8	16.00 15.49 (1, 5, 11, 22, 47)				

In order to allow for the possibility of stopping after one group, the values of h is taken to satisfy $h \le \frac{n(n+1)}{2} - k$. For example, for $n = 4$ and $k = 0$, the reference value h is taken to be smaller than or equal to 10, since $\frac{n(n+1)}{2} - k = \frac{4(4+1)}{2} - 0 = 10$.

The five percentiles are displayed in boxplot-like‡ graphs in Figure 3.2 for all the (h,k)-combinations that are shaded in Table 3.11. It clearly shows the effects of h and k on the run length distribution. Figure 3.2 describes the run-length distribution when the process is in-control. We would prefer a "boxplot" with a high valued (large) in-control average run length and a small spread. The "boxplots" are classified into 3 categories, namely, small ($h + k \le 4$),

*The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.

† See SAS Program 7 in Appendix B for the calculation of the values in Table 3.11.

‡ It should be noted that these boxplot-like graphs differ from standard box plots. In the latter case the whiskers are drawn from the ends of the box to the smallest and largest values inside specified limits, whereas, in the case of the boxplot-like graphs, the whiskers are drawn from the ends of the box to the 5^{th} and 95^{th} percentiles, respectively. In this thesis "boxplot" will refer to a boxplot-like graph from this point forward.

moderate ($5 \le h+k \le 8$) and large ($h+k \ge 9$). If the sum of the reference value, k, and the decision interval, h, is small (moderate or large), the corresponding "boxplot" is classified under small (moderate or large). For example, where $h+k=4$, the "boxplot" is classified as small, since the ARL_0^+, $SDRL$ and percentile values are small for $n=4$. In contrast, where $h+k=10$, the "boxplot" is classified as large, since the ARL_0^+, $SDRL$ and percentile values are large for $n=4$.

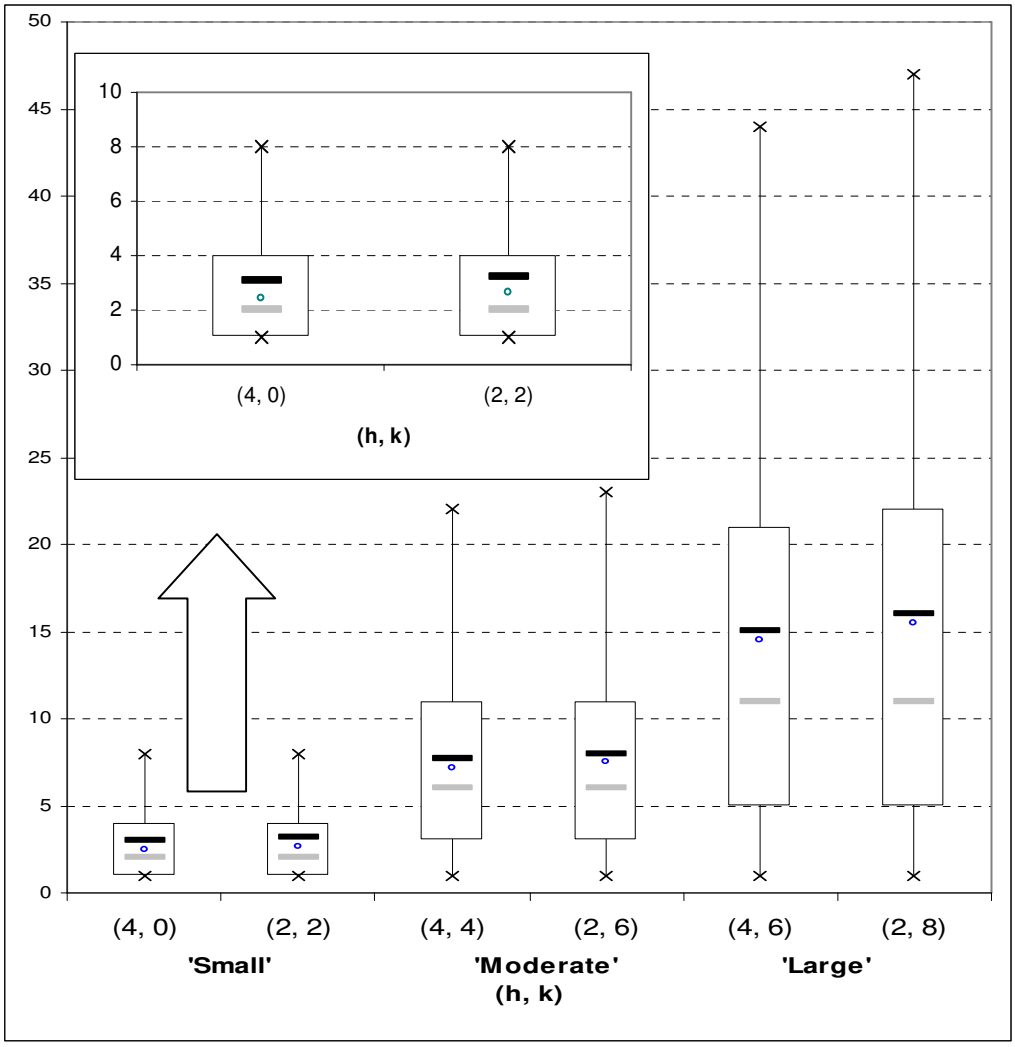

Figure 3.2. Boxplot-like graphs for the in-control run length distribution of various upper one-sided CUSUM signed-rank charts when $n=4$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "○" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Example 3.4

An upper one-sided CUSUM signed-rank chart where the sample size is odd (n=5)

The statistical properties of an upper one-sided CUSUM signed-rank chart with a decision interval of 6 $(h=6)$, a reference value of 3 $(k=3)$ and a sample size of 5 $(n=5)$ is examined. We start by examining the pmf of the well-known Wilcoxon signed-rank statistic T^+, since the plotting statistic SR_i is linearly related to T^+ (see equation (3.2)).

Table 3.12. Enumeration for the distribution of T^+ for a sample size of 5.

Value of T^+	**Ranks associated with positive differences**	**Number of sample points $u(t)$**	$P(T^+=t)$	$P(T^+\leq t)$
0		1	$1/32$	$1/32$
1	1	1	$1/32$	$2/32$
2	2	1	$1/32$	$3/32$
3	{1,2}; {3}	2	$2/32$	$5/32$
4	{1,3}; {4}	2	$2/32$	$7/32$
5	{1,4}; {2,3}; {5}	3	$3/32$	$10/32$
6	{1,2,3}; {1,5}; {2,4}	3	$3/32$	$13/32$
7	{1,2,4}; {2,5}; {3,4}	3	$3/32$	$16/32$
8	{1,2,5}; {1,3,4}; {3,5}	3	$3/32$	$19/32$
9	{1,3,5}; {2,3,4}; {4,5}	3	$3/32$	$22/32$
10	{1,2,3,4}; {1,4,5}; {2,3,5}	3	$3/32$	$25/32$
11	{1,2,3,5}; {2,4,5}	2	$2/32$	$27/32$
12	{1,2,4,5}; {3,4,5}	2	$2/32$	$29/32$
13	{1,3,4,5}	1	$1/32$	$30/32$
14	{2,3,4,5}	1	$1/32$	$31/32$
15	{1,2,3,4,5}	1	$1/32$	$32/32$

From Table 3.12 if follows that the pmf of T^+ when the sample size is 5 is

$$f_{T^+}(t) = P(T^+ = t) = \begin{cases} 1/32 & t = 0,1,2,13,14,15 \\ 2/32 & t = 3,4,11,12 \\ 3/32 & t = 5,6,7,8,9,10 \\ 0 & \text{otherwise} \end{cases}$$

The reference value was taken to be equal to three, because this leads to the sum $\sum(SR_i - k)$ being equal to even values which reduces the size of the state space for the Markov chain. For $h = 6$ we have that $\Omega^+ = \{\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3\} = \{0,2,4,6\}$ with $0 = \varsigma_0 < \varsigma_1 < \varsigma_2 < \varsigma_3 = h$. The state space is calculated using equation (3.3) and the calculations are shown in Table 3.13.

Table 3.13. Calculation of the state space when $h = 6$, $k = 3$ and $n = 5$.

SR_t	$S_{t-1}^+ + SR_t - k$	$\max\{0, S_{t-1}^+ + SR_t - k\}$	$S_t^+ = \min\{h, \max\{0, S_{t-1}^+ + SR_t - k\}\}$
-15	-18*	0	0
-13	-16	0	0
-11	-14	0	0
-9	-12	0	0
-7	-10	0	0
-5	-8	0	0
-3	-6	0	0
-1	-4	0	0
1	-2	0	0
3	0	0	0
5	2	2	2
7	4	4	4
9	6	6	6
11	8	8	6
13	10	10	6
15	12	12	6

* Note: Since only the state space needs to be described, S_{t-1}^+ can be any value from Ω^+ and we therefore take, without loss of generality, $S_{t-1}^+ = 0$. Any other possible value for S_{t-1}^+ would lead to the same Ω^+.

Table 3.14. Classification of the states.

State number	Description of the state	Absorbent (A)/ Non-absorbent (NA)
0	$S_t^+ = 0$	NA
1	$S_t^+ = 2$	NA
2	$S_t^+ = 4$	NA
3	$S_t^+ = 6$	A

From Table 3.14 we see that there are three non-absorbent states, i.e. $r = 3$, and one absorbent state, i.e. $s = 1$. Therefore, the corresponding TPM will be a $(r+s)\times(r+s) = 4\times 4$ matrix. It can be shown (see Table 3.15) that the TPM is given by

$$TPM_{4\times 4} = \begin{pmatrix} p_{00} & p_{02} & p_{04} & p_{06} \\ p_{20} & p_{22} & p_{24} & p_{26} \\ p_{40} & p_{42} & p_{44} & p_{46} \\ p_{60} & p_{62} & p_{64} & p_{66} \end{pmatrix} = \left(\begin{array}{ccccc} {}^{22}\!/_{32} & {}^{3}\!/_{32} & {}^{2}\!/_{32} & | & {}^{5}\!/_{32} \\ {}^{19}\!/_{32} & {}^{3}\!/_{32} & {}^{3}\!/_{32} & | & {}^{7}\!/_{32} \\ {}^{16}\!/_{32} & {}^{3}\!/_{32} & {}^{3}\!/_{32} & | & {}^{10}\!/_{32} \\ - & - & - & - & - \\ 0 & 0 & 0 & | & 1 \end{array}\right) = \left(\begin{array}{ccc} Q_{3\times 3} & | & \underline{p}_{3\times 1} \\ - & - & - \\ \underline{0}'_{1\times 3} & | & 1_{1\times 1} \end{array}\right)$$

where the essential transition probability sub-matrix $Q_{3\times 3} : (NA \to NA)$ is an $r \times r = 3\times 3$ matrix, $\underline{p}_{3\times 1} : (NA \to A)$ is an $(r+s-1)\times 1 = 3\times 1$ column vector, $\underline{0}'_{1\times 3} : (A \to NA)$ is a $1\times(r+s-1) = 1\times 3$ row vector and $1_{1\times 1} : (A \to A)$ represents the scalar value one. The calculation of the one-step transition probabilities are given for illustration in Table 3.15.

Recall that the probabilities in the last column of the TPM are calculated using the fact that $\sum_{j\in\Omega} p_{ij} = 1 \quad \forall i$ (see equation (2.18)). Therefore,

$p_{06} = 1 - (p_{00} + p_{02} + p_{04}) = 1 - ({}^{22}\!/_{32} + {}^{3}\!/_{32} + {}^{2}\!/_{32}) = {}^{5}\!/_{32}$;

$p_{26} = 1 - (p_{20} + p_{22} + p_{24}) = 1 - ({}^{19}\!/_{32} + {}^{3}\!/_{32} + {}^{3}\!/_{32}) = {}^{7}\!/_{32}$;

$p_{46} = 1 - (p_{40} + p_{42} + p_{44}) = 1 - ({}^{16}\!/_{32} + {}^{3}\!/_{32} + {}^{3}\!/_{32}) = {}^{10}\!/_{32}$;

$p_{66} = 1 - (p_{60} + p_{62} + p_{64}) = 1 - (0 + 0 + 0) = 1$.

Table 3.15. The calculation of the transition probabilities when $h=6$, $k=3$ and $n=5$.

p_{00} $= P(S_t = 0 \mid S_{t-1} = 0)$ $= P(\min\{6, \max\{0,0+SR_t-3\}\}=0)$ $= P(\max\{0, SR_t-3\}=0)$ $= P(SR_t-3\le 0)$ $= P(SR_t \le 3)$ $= P(2T^+-15\le 3)$ $= P(T^+\le 9)$ $= \frac{22}{32}$	p_{02} $= P(S_t = 2 \mid S_{t-1} = 0)$ $= P(\min\{6, \max\{0,0+SR_t-3\}\}=2)$ $= P(\max\{0, SR_t-3\}=2)$ $= P(SR_t-3=2)$ $= P(SR_t=5)$ $= P(2T^+-15=5)$ $= P(T^+=10)$ $= \frac{3}{32}$	p_{02} $= P(S_t = 4 \mid S_{t-1} = 0)$ $= P(\min\{6, \max\{0,0+SR_t-3\}\}=4)$ $= P(\max\{0, SR_t-3\}=4)$ $= P(SR_t-3=4)$ $= P(SR_t=7)$ $= P(2T^+-15=7)$ $= P(T^+=11)$ $= \frac{2}{32}$
p_{20} $= P(S_t = 0 \mid S_{t-1} = 2)$ $= P(\min\{6, \max\{0,2+SR_t-3\}\}=0)$ $= P(\max\{0, SR_t-1\}=0)$ $= P(SR_t-1\le 0)$ $= P(SR_t \le 1)$ $= P(2T^+-15\le 1)$ $= P(T^+\le 8)$ $= \frac{19}{32}$	p_{22} $= P(S_t = 2 \mid S_{t-1} = 2)$ $= P(\min\{6, \max\{0,2+SR_t-3\}\}=2)$ $= P(\max\{0, SR_t-1\}=2)$ $= P(SR_t-1=2)$ $= P(SR_t=3)$ $= P(2T^+-15=3)$ $= P(T^+=9)$ $= \frac{3}{32}$	p_{24} $= P(S_t = 4 \mid S_{t-1} = 2)$ $= P(\min\{6, \max\{0,2+SR_t-3\}\}=4)$ $= P(\max\{0, SR_t-1\}=4)$ $= P(SR_t-1=4)$ $= P(SR_t=5)$ $= P(2T^+-15=5)$ $= P(T^+=10)$ $= \frac{3}{32}$
p_{40} $= P(S_t = 0 \mid S_{t-1} = 4)$ $= P(\min\{6, \max\{0,4+SR_t-3\}\}=0)$ $= P(\max\{0, SR_t+1\}=0)$ $= P(SR_t+1\le 0)$ $= P(SR_t \le -1)$ $= P(2T^+-15\le -1)$ $= P(T^+\le 7)$ $= \frac{16}{32}$	p_{42} $= P(S_t = 2 \mid S_{t-1} = 4)$ $= P(\min\{6, \max\{0,4+SR_t-3\}\}=2)$ $= P(\max\{0, SR_t+1\}=2)$ $= P(SR_t+1=2)$ $= P(SR_t=1)$ $= P(2T^+-15=1)$ $= P(T^+=8)$ $= \frac{3}{32}$	p_{44} $= P(S_t = 4 \mid S_{t-1} = 4)$ $= P(\min\{6, \max\{0,4+SR_t-3\}\}=4)$ $= P(\max\{0, SR_t+1\}=4)$ $= P(SR_t+1=4)$ $= P(SR_t=3)$ $= P(2T^+-15=3)$ $= P(T^+=9)$ $= \frac{3}{32}$
p_{60} $= P(S_t = 0 \mid S_{t-1} = 6)$ $= 0$ *	p_{62} $= P(S_t = 2 \mid S_{t-1} = 6)$ $= 0$	p_{64} $= P(S_t = 4 \mid S_{t-1} = 6)$ $= 0$

* The probability equals zero, because it is impossible to go from an absorbent state to a non-absorbent state.

The formulas of the moments and some characteristics of the run length distribution have been studied by Fu, Spiring and Xie (2002) and Fu and Lou (2003) – see equations (2.41) to (2.45). By substituting $\underline{\xi}_{1\times 3} = (1 \quad 0 \quad 0)$, $Q_{3\times 3} = \begin{pmatrix} 22/32 & 3/32 & 2/32 \\ 19/32 & 3/32 & 3/32 \\ 16/32 & 3/32 & 3/32 \end{pmatrix}$ and $\underline{1}_{3\times 1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ into these equations, we obtain the following:

$$ARL = E(N) = \underline{\xi}(I-Q)^{-1}\underline{1} = 5.79$$

$$E(N^2) = \underline{\xi}(I+Q)(I-Q)^{-2}\underline{1} = 60.14$$

$$SDRL = \sqrt{Var(N)} = \sqrt{E(N^2) - (E(N))^2} = 5.16$$

5^{th} percentile $= p_5 = 1$

25^{th} percentile $= p_{25} = 2$

Median $= 50^{th}$ percentile $= p_{50} = 4$

75^{th} percentile $= p_{75} = 8$

95^{th} percentile $= p_{95} = 16$

Other values of *h*, *k* and *n* were also considered and the results are given in Table 3.16.

Table 3.16. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for the upper one-sided CUSUM signed-rank chart when $n = 5$†.

k	h						
	2	**4**	**6**	**8**	**10**	**12**	**14**
1	2.46 1.90 (1, 1, 2, 3, 6)	3.11 2.53 (1, 1, 2, 4, 8)	4.08 3.42 (1, 2, 3, 5, 11)	5.14 4.38 (1, 2, 4, 7, 14)	6.71 5.73 (1, 3, 5, 9, 18)	8.29 7.13 (1, 3, 6, 11, 22)	10.46 8.99 (2, 4, 8, 14, 28)
3	3.20 2.65 (1, 1,2, 4, 8)	4.39 3.82 (1, 2, 3, 6, 12)	5.79 5.16 (1, 2, 4, 8, 16)	8.13 7.34 (1, 3, 6, 11, 23)	10.68 9.75 (1, 4, 8, 14, 30)	14.78 13.56 (2, 5, 11, 20, 42)	
5	4.57 4.04 (1, 2, 3, 6, 13)	6.24 5.69 (1, 2, 4, 8, 18)	9.44 8.79 (1, 3, 7, 13, 27)	13.04 12.31 (1, 4, 9, 18, 38)	20.16 19.22 (2, 6, 14, 28, 59)		
7	6.40 5.88 (1, 2, 5, 9, 18)	10.24 9.68 (1, 3, 7, 14, 30)	14.77 14.18 (1, 5, 10, 20, 43)	25.17 24.43 (2, 8, 18, 35, 74)			
9	10.67 10.15 (1, 3, 8, 15, 31)	15.75 15.22 (1, 5, 11, 22, 46)	29.15 28.55 (2, 9, 20, 40, 86)				
11	16.00 15.49 (1, 5, 11, 22, 47)	31.03 30.50 (2, 9, 22, 43, 92)					
13	32.00 31.50 (2, 10, 22, 44, 95)						

* The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.

† See SAS program 7 in Appendix B for the calculation of the values in Table 3.16.

The five percentiles are displayed in boxplot-like graphs in Figure 3.3 for all the (h,k)-combinations that are shaded in Table 3.16. It clearly shows the effects of h and k on the run length distribution. Figure 3.3 describes the run-length distribution when the process is in-control. We would prefer a "boxplot" with a high valued (large) in-control average run length and a small spread. The "boxplots" are classified into 3 categories, namely small ($h+k \leq 5$), moderate ($6 \leq h+k \leq 10$) and large ($h+k \geq 11$).

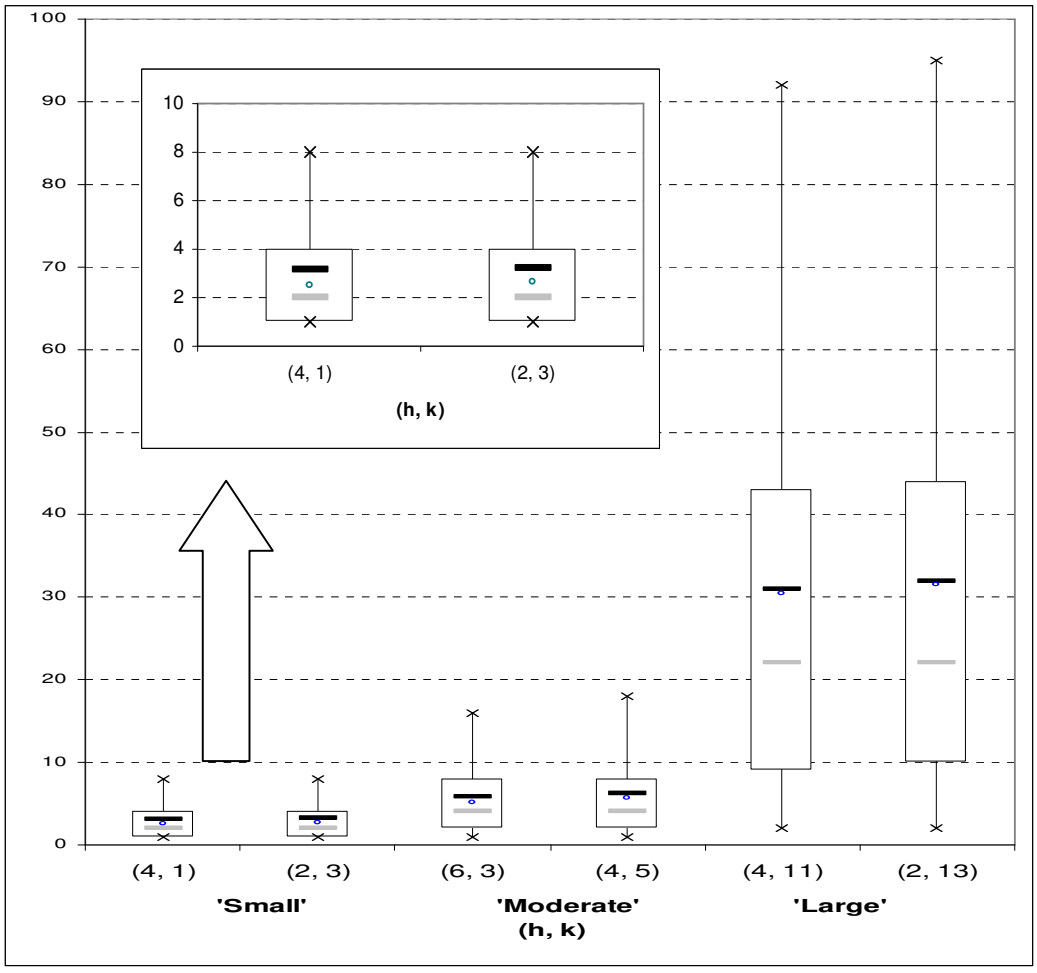

Figure 3.3. Boxplot-like graphs for the in-control run length distribution of various upper one-sided CUSUM signed-rank charts when $n=5$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL** and *MRL*, respectively.

* For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Examples 3.3 and 3.4 illustrated the Markov chain approach used to calculate run length characteristics for n even and odd, respectively. On the performance side, note that the largest in-control average run length that the upper one-sided CUSUM signed-rank can obtain is 2^n. Therefore, for a sample size of 4 the largest ARL_0^+ equals $2^4 = 16$ (this is obtained when $h = 2$ and $k = 8$). Thus, a large number of false alarms will be signaled by this chart leading to a possible loss of time and resources. Compared to this, for a sample of size 5 the largest ARL_0^+ equals $2^5 = 32$ (this is obtained when $h = 2$ and $k = 13$). Both examples considered sample sizes that may be considered "small". Some results will be given for larger sample sizes ($n = 6$ and 10).

Table 3.17. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for the upper one-sided CUSUM signed-rank chart when $n = 6$†.

k	h								
	2	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**
1	2.37 1.80 (1, 1, 2, 3, 6)	2.86 2.28 (1, 1, 2, 4, 7)	3.39 2.79 (1, 1, 3, 4, 9)	4.08 3.42 (1, 2, 3, 5, 11)	5.03 4.25 (1, 2, 4, 7, 13)	6.10 5.17 (1, 2, 5, 8, 16)	7.24 6.17 (1, 3, 5, 10, 20)	8.72 7.41 (2, 3, 6, 12, 23)	10.21 8.69 (2, 4, 8, 14, 28)
3	2.91 2.36 (1, 1, 2, 4, 8)	3.51 2.95 (1, 1, 3, 5, 9)	4.33 3.74 (1, 2, 3, 6, 12)	5.55 4.87 (1, 2, 4, 7, 15)	7.00 6.21 (1, 3, 5, 9, 19)	8.63 7.72 (1, 3, 6, 12, 24)	10.99 9.86 (2, 4, 8, 15, 31)	13.43 12.12 (2, 5, 10, 18, 38)	16.78 15.18 (2, 6, 12, 23, 47)
5	3.56 3.01 (1, 1, 3, 5, 10)	4.49 3.94 (1, 2, 3, 6, 12)	5.95 5.35 (1, 2, 4, 8, 17)	7.82 7.14 (1, 3, 6, 11, 22)	10.02 9.25 (1, 3, 7, 14, 28)	13.55 12.60 (2, 5, 10, 18, 39)	17.39 16.29 (2, 6, 12, 24, 50)	23.44 22.07 (2, 8, 17, 32, 67)	
7	4.57 4.04 (1, 2, 3, 6, 13)	6.24 5.70 (1, 2, 4, 8, 18)	8.50 7.91 (1, 3, 6, 12, 24)	11.26 10.61 (1, 4, 8, 15, 32	16.17 15.39 (2, 5, 11, 22, 47)	21.79 20.90 (2, 7, 15, 30, 63)	32.01 30.88 (3, 10, 23, 44, 94)		
9	6.40 5.88 (1, 2, 5, 9, 18)	8.96 8.42 (1, 3, 6, 12, 26)	12.16 11.60 (1, 4, 9, 17, 35)	18.48 17.83 (2, 6, 13, 25, 54)	25.89 25.17 (2, 8, 18, 36, 76)	41.56 40.64 (2, 13, 29, 57, 123)			
11	9.14 8.63 (1, 3, 6, 12, 26)	12.64 12.12 (1, 4, 9, 17, 37)	20.05 19.48 (2, 6, 14, 28, 59)	28.88 28.27 (2, 9, 20, 40, 85)	50.26 49.52 (3, 15, 35, 69, 149)				
13	12.80 12.29 (1, 4, 9, 18, 37)	20.90 20.37 (2, 6, 15, 29, 62)	30.76 30.22 (2, 9, 21, 42, 91)	56.62 55.99 (3, 17, 39, 78, 168)					
15	21.33 20.83 (2, 6, 15, 29, 63)	31.75 31.24 (2, 9, 22, 44, 94)	61.08 60.53 (4, 18, 43, 84, 182)						
17	32.00 31.50 (2, 10, 22, 44, 95)	63.02 62.50 (4, 18, 44, 87, 188)							
19	64.00 63.50 (4, 19, 45, 89, 191)								

* The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.

† See SAS Program 7 in Appendix B for the calculation of the values in Table 3.17.

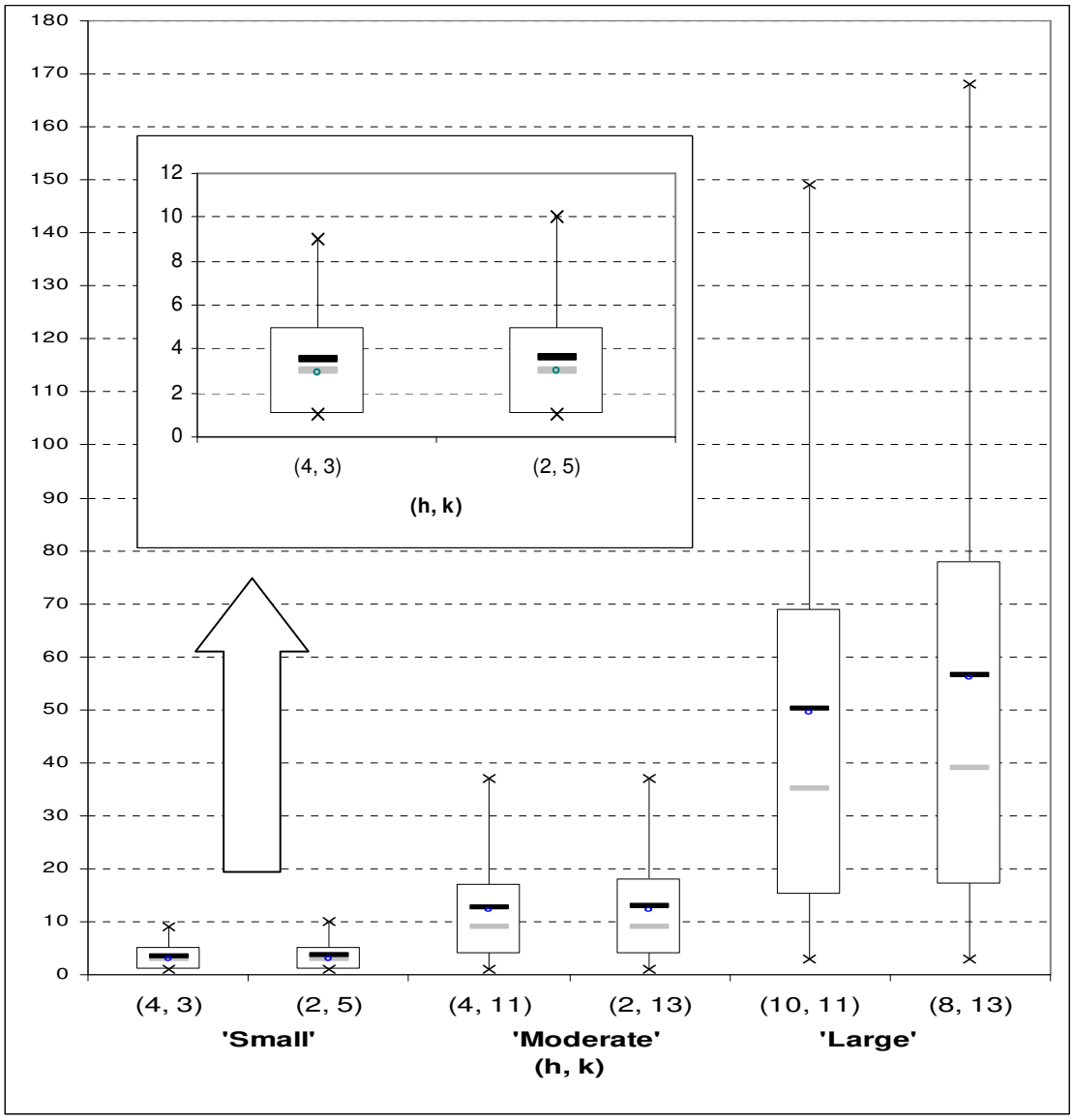

Figure 3.4. Boxplot-like graphs for the in-control run length distribution of various upper one-sided CUSUM signed-rank charts when $n = 6$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL*[*] and *MRL*, respectively[†].

[*] For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

[†] The "boxplots" are classified into 3 categories, namely small ($h+k \leq 7$), moderate ($8 \leq h+k \leq 16$) and large ($h+k \geq 17$).

Table 3.18. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for samples of size $n = 10$ for $h = 2,4,...,14$ and $k = 1,3,...,23$ for the upper one-sided CUSUM signed-rank chart†.

k	h						
	2	**4**	**6**	**8**	**10**	**12**	**14**
1	2.17 1.59 (1, 1, 2, 3, 5)	2.36 1.78 (1, 1, 2, 3, 6)	2.57 1.99 (1, 1, 2, 3, 7)	2.81 2.22 (1, 1, 2, 4, 7)	3.07 2.47 (1, 1, 2, 4, 8)	3.36 2.73 (1, 1, 3, 4, 9)	3.68 3.02 (1, 2, 3, 5, 10)
3	2.36 1.80 (1, 1, 2, 3, 6)	2.59 2.02 (1, 1, 2, 3, 7)	2.84 2.27 (1, 1, 2, 4, 7)	3.12 2.54 (1, 1, 2, 4, 8)	3.44 2.84 (1, 1, 3, 5, 9)	3.79 3.17 (1, 2, 3, 5, 10)	4.18 3.52 (1, 2, 3, 6, 11)
5	2.60 2.04 (1, 1, 2, 3, 7)	2.87 2.31 (1, 1, 2, 4, 7)	3.16 2.60 (1, 1, 2, 4, 8)	3.50 2.93 (1, 1, 3, 5, 9)	3.88 3.29 (1, 2, 3, 5, 10)	4.31 3.69 (1, 2, 3, 6, 12)	4.79 4.14 (1, 2, 4, 6, 13)
7	2.88 2.32 (1, 1, 2, 4, 8)	3.19 2.64 (1, 1, 2, 4, 8)	3.55 2.99 (1, 1, 3, 5, 10)	3.96 3.39 (1, 2, 3, 5, 11)	4.42 3.83 (1, 2, 3, 6, 12)	4.95 4.34 (1, 2, 4, 7, 14)	5.55 4.91 (1, 2, 4, 7, 15)
9	3.20 2.65 (1, 1, 2, 4, 8)	3.58 3.03 (1, 1, 3, 5, 10)	4.01 3.46 (1, 2, 3, 5, 11)	4.51 3.94 (1, 2, 3, 6, 12)	5.08 4.50 (1, 2, 4, 7, 14)	5.75 5.14 (1, 2, 4, 8, 16)	6.49 5.85 (1, 2, 5, 9, 18)
11	3.59 3.05 (1, 1, 3, 5, 10)	4.05 3.51 (1, 2, 3, 5, 11)	4.57 4.03 (1, 2, 3, 6, 13)	5.19 4.63 (1, 2, 4, 7, 14)	5.91 5.33 (1, 2, 4, 8, 17)	6.73 6.12 (1, 2, 5, 9, 19)	7.69 7.05 (1, 3, 6, 10, 22)
13	4.06 3.53 (1, 2, 3, 5, 11)	4.61 4.08 (1, 2, 3, 6, 13)	5.26 4.72 (1, 2, 4, 7, 15)	6.03 5.48 (1, 2, 4, 8, 17)	6.92 6.35 (1, 2, 5, 9, 20)	7.97 7.37 (1, 3, 6, 11, 23)	9.24 8.60 (1, 3, 7, 13, 26)
15	4.63 4.10 (1, 2, 3, 6, 13)	5.31 4.78 (1, 2, 4, 7, 15)	6.12 5.59 (1, 2, 4, 8, 17)	7.06 6.52 (1, 2, 5, 10, 20)	8.20 7.63 (1, 3, 6, 11, 23)	9.58 8.99 (1, 3, 7, 13, 28)	11.19 10.56 (1, 4, 8, 15, 32)
17	5.33 4.81 (1, 2, 4, 7, 15)	6.18 5.65 (1, 2, 4, 8, 17)	7.17 6.64 (1, 2, 5, 10, 20)	8.37 7.83 (1, 3, 6, 11, 24)	9.86 9.30 (1, 3, 7, 13, 28)	11.60 11.02 (1, 4, 8, 16, 34)	13.74 13.12 (1, 4, 10, 19, 40)
19	6.21 5.68 (1, 2, 4, 8, 18)	7.23 6.71 (1, 2, 5, 10, 21)	8.50 7.97 (1, 3, 6, 12, 24)	10.07 9.53 (1, 3, 7, 14, 29)	11.93 11.37 (1, 4, 8, 16, 35)	14.24 13.66 (1, 5, 10, 20, 42)	17.12 16.50 (1, 5, 12, 23, 50)
21	7.26 6.74 (1, 2, 5, 10, 21)	8.57 8.05 (1, 3, 6, 12, 25)	10.21 9.69 (1, 3, 7, 14, 30)	12.18 11.64 (1, 4, 9, 17, 35)	14.64 14.08 (1, 5, 10, 20, 43)	17.73 17.15 (1, 6, 12, 24, 52)	21.60 21.00 (2, 7, 15, 30, 64)
23	8.61 8.09 (1, 3, 6, 12, 25)	10.30 9.79 (1, 3, 7, 14, 30)	12.34 11.82 (1, 4, 9, 17, 36)	14.93 14.39 (1, 5, 11, 20, 44)	18.20 17.65 (1, 6, 13, 25, 53)	22.36 21.79 (2, 7, 16, 31, 66)	28.16 27.56 (2, 9, 20, 39, 83)

* The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.

† See SAS Program 7 in Appendix B for the calculation of the values in Table 3.18.

Table 3.18 continued for $h = 2, 4, ..., 14$ and $k = 25, 27, ..., 53$.

k	h						
	2	**4**	**6**	**8**	**10**	**12**	**14**
25	10.34 9.83 (1, 3, 7, 14, 30)	12.44 11.93 (1, 4, 9, 17, 36)	15.12 14.60 (1, 5, 11, 21, 44)	18.55 18.02 (1, 6, 13, 26, 55)	22.93 22.38 (2, 7, 16, 32, 68)	29.14 28.57 (2, 9, 20, 40, 86)	37.30 36.70 (2, 11, 26, 51, 111)
27	12.49 11.98 (1, 4, 9, 17, 36)	15.23 14.72 (1, 5, 11, 21, 45)	18.77 18.25 (1, 6, 13, 26, 55)	23.33 22.80 (2, 7, 16, 32, 69)	29.87 29.33 (2, 9, 21, 41, 88)	38.56 38.00 (3, 11, 27, 53, 114)	49.52 48.94 (3, 15, 35, 68, 147)
29	15.28 14.78 (1, 5, 11, 21, 45)	18.91 18.40 (1, 6, 13, 26, 56)	23.59 23.08 (2, 7, 17, 33, 70)	30.39 29.87 (2, 9, 21, 42, 90)	39.50 38.96 (3, 12, 28, 55, 117)	51.09 50.53 (3, 15, 36, 71, 152)	67.68 67.10 (4, 20, 47, 94, 202)
31	18.96 18.46 (1, 6, 13, 26, 56)	23.75 23.24 (2, 7, 17, 33, 70)	30.74 30.22 (2, 9, 21, 42, 91)	40.17 39.64 (3, 12, 28, 55, 119)	52.23 51.70 (3, 15, 36, 72, 155)	69.70 69.14 (4, 20, 48, 96, 208)	95.33 94.76 (5, 28, 66, 132, 284)
33	23.81 23.31 (2, 7, 17, 33, 70)	30.94 30.43 (2, 9, 22, 43, 92)	40.60 40.09 (3, 12, 28, 56, 121)	53.02 52.51 (3, 16, 37, 73, 158)	71.14 70.61 (4, 21, 49, 98, 212)	98.00 97.46 (6, 9, 68, 136, 292)	137.20 136.63 (8, 40, 95, 190, 410)
35	31.03 30.53 (2, 9, 22, 43, 92)	40.86 40.35 (3, 12, 28, 56, 121)	53.53 53.02 (3, 16, 37, 74, 159)	72.12 71.61 (4, 21, 50, 100, 215)	99.90 99.37 (6, 29, 69, 138, 298)	140.75 140.21 (8, 41, 98, 195, 421)	194.51 193.96 (11, 56, 135, 269, 582)
37	40.96 40.46 (3, 12, 29, 57, 122)	53.80 53.29 (3, 16, 37, 74, 160)	72.71 72.20 (4, 21, 51, 101, 217)	101.12 100.60 (6, 29, 70, 140, 302)	143.15 142.63 (8, 42, 99, 198,428)	198.67 198.14 (11, 58, 138, 275, 594)	323.14 322.58 (17, 93, 224, 448, 967)
39	53.89 53.39 (3, 16, 38, 75, 160)	73.02 72.51 (4, 21, 51, 101, 218)	101.84 101.33 (6, 30, 71, 141, 304)	144.68 144.17 (8, 42, 100, 200, 432)	201.42 200.90 (11, 58, 140, 279, 602)	330.31 329.78 (17, 95, 229, 458, 988)	490.25 489.71 (26, 141, 340, 679, 1468)
41	73.14 72.64 (4, 21, 51, 101, 218)	102.24 101,74 (6, 30, 71, 142, 305)	145.61 145.11 (8, 42, 101, 202, 435)	203.16 203.65 (11, 59, 141, 281, 608)	335.17 334.65 (18, 97, 232, 464, 1003)	499.40 498.88 (26, 144, 346, 692, 1495)	973.74 973.19 (50, 281, 675, 1350, 2916)
43	102.40 101.90 (6, 30, 71, 142, 306)	146.10 145.60 (8, 42, 101, 202, 437)	204.16 203.66 (11, 59, 142, 283, 611)	338.24 337.73 (18, 98, 235, 469, 1012)	505.29 504.77 (26, 146, 350, 700, 1513)	994.57 994.05 (52, 286, 690, 1379, 2978)	
45	146.29 145.78 (8, 42, 102, 203, 437)	204.64 204.14 (11, 59, 142, 283, 612)	340.00 339.50 (18, 98, 236, 471, 1018)	508.76 508.25 (27, 147, 353, 705, 1523)	1008.16 1007.64 (52, 290, 699, 1397, 3019)		
47	204.80 204.30 (11, 59, 142, 284, 613)	340.89 340.39 (18, 98, 236, 472, 1020)	510.75 510.25 (27, 147, 354, 708, 1529)	1016.04 1015.53 (53, 293, 704, 1408, 3043)			
49	341.33 340.83 (18, 99, 237, 473, 1022)	511.75 511.25 (27, 148, 355, 709, 1532)	1021.00 1020.50 (53, 294, 708, 1415, 3058)				
51	512.00 511.50 (27, 148, 355, 710, 1533)	1023.00 1022.50 (53, 295, 709, 1418, 3064)					
53	1024.00 1023.50 (53, 295, 710, 1419, 3067)						

Table 3.18 continued for h = 16, 18, ..., 28 and k = 1, 3, ..., 25.

k	***h***						
	16	**18**	**20**	**22**	**24**	**26**	**28**
1	4.03 3.33 (1, 2, 3, 5, 11)	4.41 3.67 (1, 2, 3, 6, 12)	4.83 4.03 (1, 2, 4, 6, 13)	5.27 4.41 (1, 2, 4, 7, 14)	5.76 4.82 (1, 2, 4, 8, 15)	6.28 5.26 (1, 3, 5, 8, 17)	6.83 5.72 (1, 3, 5, 9, 18)
3	4.61 3.91 (1, 2, 3, 6, 12)	5.09 4.34 (1, 2, 4, 7, 14)	5.60 4.80 (1, 2, 4, 7, 15)	6.18 5.31 (1, 2, 5, 8, 17)	6.80 5.86 (1, 3, 5, 9, 18)	7.47 6.44 (1, 3, 6, 10, 20)	8.19 7.07 (1, 3, 6, 11, 22)
5	5.33 4.64 (1, 2, 4, 7, 15)	5.93 5.18 (1, 2, 4, 8, 16)	6.59 5.79 (1, 2, 5, 9, 18)	7.34 6.47 (1, 3, 5, 10, 20)	8.14 7.19 (1, 3, 6, 11, 22)	9.03 7.99 (1, 3, 7, 12, 25)	10.00 8.87 (1, 4, 7, 13, 28)
7	6.23 5.54 (1, 2, 5, 8, 17)	6.99 6.25 (1, 3, 5, 9, 19)	7.86 7.06 (1, 3, 6, 11, 22)	8.82 7.95 (1, 3, 6, 12, 25)	9.90 8.95 (1, 4, 7, 13, 28)	11.10 10.06 (1, 4, 8, 15, 31)	12.43 11.30 (2, 4, 9, 17, 35)
9	7.36 6.68 (1, 3, 5, 10, 21)	8.37 7.63 (1, 3, 6, 11, 24)	9.49 8.69 (1, 3, 7, 13, 27)	10.77 9.90 (1, 4, 8, 15, 31)	12.23 11.28 (1, 4, 9, 17, 35)	13.88 12.85 (2, 5, 10, 19, 40)	15.83 14.69 (2, 5, 11, 22, 45)
11	8.83 8.15 (1, 3, 6, 12, 25)	10.12 9.38 (1, 3, 7, 14, 29)	11.62 10.83 (1, 4, 8, 16, 33)	13.36 12.50 (1, 4, 10, 18, 38)	15.38 14.44 (2, 5, 11, 21, 44)	17.82 16.78 (2, 6, 13, 24, 51)	20.61 19.47 (2, 7, 15, 28, 59)
13	10.69 10.01 (1, 4, 8, 15, 31)	12.41 11.69 (1, 4, 9, 17, 36)	14.45 13.67 (1, 5, 10, 20, 42)	16.85 16.00 (2, 5, 12, 23, 49)	19.85 18.91 (2, 6, 14, 27, 58)	23.36 22.32 (2, 7, 17, 32, 68)	27.32 26.19 (2, 9, 19, 37, 80)
15	13.13 12.46 (1, 4, 9, 18, 38)	15.46 14.74 (1, 5, 11, 21, 45)	18.26 17.48 (2, 6, 13, 25, 53)	21.85 21.00 (2, 7, 15, 30, 64)	26.15 25.21 (2, 8, 18, 36, 76)	31.11 30.09 (3, 10, 22, 43, 91)	37.23 36.10 (3, 12, 26, 51, 109)
17	16.36 15.70 (1, 5, 12, 22, 48)	19.54 18.84 (2, 6, 14, 27, 57)	23.74 22.97 (2, 7, 17, 33, 70)	28.88 28.03 (2, 9, 20, 40, 85)	34.92 34.00 (3, 11, 24, 48, 103)	42.59 41.57 (3, 13, 30, 59, 126)	52.23 51.12 (4, 16, 37, 72, 154)
19	20.67 20.02 (2, 6, 15, 28, 61)	25.45 24.74 (2, 8, 18, 35, 75)	31.43 30.67 (2, 10, 22, 43, 93)	38.60 37.77 (3, 12, 27, 53, 114)	47.93 47.02 (3, 14, 34, 66, 142)	60.03 59.03 (4, 18, 42, 83, 178)	75.33 74.23 (5, 22, 53, 104, 223)
21	26.93 26.28 (2, 8, 19, 37, 79)	33.72 33.02 (2, 10, 24, 46, 100)	41.99 41.23 (3, 13, 29, 58, 124)	53.02 52.20 (3, 16, 37, 73, 157)	67.73 66.84 (4, 20, 47, 94, 201)	86.89 85.90 (5, 26, 61, 120, 258)	110.04 108.97 (7, 32, 77, 152, 327)
23	35.69 35.04 (2, 11, 25, 49, 106)	44.98 44.29 (3, 13, 31, 62, 133)	57.66 57.66 (4, 17, 40, 80, 171)	75.02 74.21 (5, 22, 52, 104, 223)	98.22 97.34 (6, 29, 68, 136, 292)	126.80 125.85 (7, 37, 88, 175, 378)	175.50 174.41 (10, 51, 122, 243, 524)
25	47.50 46.87 (3, 14, 33, 66, 141)	61.69 61.02 (4, 18, 43, 85, 183)	81.56 80.84 (5, 24, 57, 113, 243)	108.78 107.99 (6, 32, 76, 150, 324)	142.88 142.04 (8, 42, 99, 198, 426)	204.49 203.53 (11, 60, 142, 283, 611)	275.50 274.45 (15, 80, 191, 382, 823)

Table 3.18 continued for h = 16, 18, ..., 28 and k = 27, 29, ..., 53.

k	***h***						
	16	**18**	**20**	**22**	**24**	**26**	**28**
27	65.03 64.41 (4, 19, 45, 90, 194)	87.16 86.50 (5, 26, 61, 121, 260)	118.11 117.40 (7, 34, 82, 163, 352)	157.47 156.72 (9, 46, 109, 218, 470)	232.48 231.62 (13, 67, 161, 322, 695)	320.44 319.51 (17, 93, 222, 444, 958)	500.06 498.96 (27, 145, 347, 693, 1496)
29	91.75 91.14 (5, 27, 64, 127, 274)	126.00 125.35 (7, 37, 88, 174, 376)	170.09 169.40 (9, 49, 118, 236, 508)	258.15 257.38 (14, 75, 179, 358, 772)	363.03 362.21 (19, 105, 252, 503, 1086)	594.47 593.50 (31, 172, 412, 824, 1779)	
31	132.34 131.74 (7, 38, 92, 183, 395)	180.44 179.81 (10, 52, 125, 250, 539)	280.37 279.68 (15, 81, 195, 388, 839)	400.92 400.19 (21, 116, 278, 556, 1200)	686.75 685.89 (36, 198, 476, 952, 2056)		
33	188.52 187.93 (10, 55, 131, 261, 564)	298.55 297.92 (16, 86, 207, 414, 893)	432.84 432.18 (23, 125, 300, 600, 1295)	770.90 770.13 (40, 222, 535, 1068, 2308)			
35	312.77 312.18 (17, 90, 217, 433, 936)	458.19 457.58 (24, 132, 318, 635, 1371)	824.79 842.11 (44, 243, 584, 1168, 2523)				
37	476.93 476.36 (25, 138, 331, 661, 1428)	899.99 899.36 (47, 259, 624, 1247, 2695)					
39	942.81 942.23 (49, 272, 654, 1307, 2823)						
41							
43							
45							
47							
49							
51							
53							

Table 3.18 continued for h = 30, 32, ..., 42 and k = 1, 3, ..., 53.

k	h						
	30	**32**	**34**	**36**	**38**	**40**	**42**
1	7.41 6.21 (1, 3, 6, 10, 20)	8.03 6.72 (1, 3, 6, 11, 21)	8.68 7.26 (2, 4, 7, 12, 23)	9.38 7.84 (2, 4, 7, 12, 25)	10.10 8.44 (2, 4, 8, 13, 27)	10.84 9.05 (2, 4, 8, 14, 29)	11.61 9.70 (2, 5, 9, 15, 31)
3	8.79 7.75 (1, 3, 7, 12, 24)	9.80 8.47 (2, 4, 7, 13, 27)	10.71 9.26 (2, 4, 8, 14, 29)	11.66 10.09 (2, 5, 9, 16, 32)	12.65 10.96 (2, 5, 9, 17, 34)	13.71 11.89 (2, 5, 10, 18, 37)	14.82 12.87 (2, 6, 11, 20, 40)
5	11.05 9.82 (2, 4, 8, 15, 31)	12.24 10.89 (2, 5, 9, 16, 34)	13.50 12.04 (2, 5, 10, 18, 37)	14.84 13.25 (2, 5, 11, 20, 41)	16.28 14.57 (2, 6, 12, 22, 45)	17.84 15.99 (3, 7, 13, 24, 50)	19.49 17.50 (3, 7, 14, 26, 54)
7	13.96 12.71 (2, 5, 10, 19, 39)	15.63 14.26 (2, 6, 11, 21, 44)	17.41 15.93 (2, 6, 13, 24, 49)	19.39 17.79 (2, 7, 14, 26, 55)	21.57 19.83 (3, 7, 16, 29, 61)	23.92 22.04 (3, 8, 17, 32, 68)	26.43 24.42 (3, 9, 19, 36, 75)
9	18.01 16.76 (2, 6, 13, 24, 51)	20.38 19.01 (2, 7, 15, 28, 58)	23.07 21.58 (3, 8, 16, 31, 66)	26.10 24.48 (3, 9, 19, 36, 75)	29.44 27.69 (3, 10, 21, 40, 85)	33.06 31.17 (3, 11, 24, 45, 95)	37.30 35.26 (4, 12, 27, 51, 108)
11	23.71 22.46 (2, 8, 17, 32, 69)	27.32 25.95 (3, 9, 19, 37, 79)	31.48 29.99 (3, 10, 22, 43, 91)	36.20 34.57 (3, 12, 26, 50, 105)	41.38 39.62 (4, 13, 29, 57, 120)	47.70 45.79 (4, 15, 34, 65, 139)	54.39 52.34 (5, 17, 38, 75, 159)
13	32.08 30.83 (3, 10, 23, 44, 94)	37.72 36.35 (3, 12, 27, 52, 110)	44.27 42.77 (4, 14, 31, 61, 130)	51.58 49.97 (4, 16, 36, 71, 151)	60.93 59.16 (5, 19, 43, 84, 179)	70.90 69.00 (5, 22, 50, 98, 209)	83.34 81.28 (6, 25, 58, 115, 246)
15	44.70 43.46 (3, 14, 31, 61, 131)	53.62 52.26 (4, 16, 38, 74, 158)	63.77 62.30 (5, 19, 45, 88, 188)	77.41 75.78 (6, 23, 54, 107, 229)	92.09 90.34 (6, 28, 64, 127, 272)	111.35 109.43 (8, 33, 78, 154, 330)	
17	64.08 62.85 (4, 19, 45, 88, 190)	77.84 76.51 (5, 23, 54, 107, 231)	97.39 95.89 (6, 29, 68, 134, 289)	118.64 117.02 (8, 35, 83, 164, 352)	148.17 146.38 (9, 44, 103, 205, 440)		
19	93.47 92.27 (6, 28, 65, 129, 278)	120.80 119.44 (7, 36, 84, 167, 359)	150.87 149.40 (9, 44, 105, 209, 449)	195.52 193.87 (12, 57, 136, 270, 582)			
21	147.14 145.92 (9, 43, 102, 204, 438)	188.54 187.22 (11, 55, 131, 261, 562)	254.71 253.20 (14, 74 , 177, 353, 760)				
23	230.72 229.53 (13, 67, 160, 319, 689)	326.23 324.86 (18, 95, 227, 452, 975)					
25	409.08 407.84 (22, 119, 284, 567, 1223)						
27							
29							
⋮							
53							

Table 3.18 continued for $h = 44, 46, ..., 54$ and $k = 1, 3, ..., 53$.

k	h					
	44	**46**	**48**	**50**	**52**	**54**
1	12.42 10.37 (2, 5, 9, 16, 33)	13.25 11.06 (2, 5, 10, 18, 35)	14.09 11.78 (3, 6, 11, 19, 38)	14.99 12.53 (3, 6, 11, 20, 40)	15.90 13.30 (3, 7, 12, 21, 42)	16.85 14.10 (3, 7, 13, 22, 45)
3	15.99 13.90 (3, 6, 12, 21, 44)	17.19 14.96 (3, 7, 13, 23, 47)	18.49 16.11 (3, 7, 14, 25, 51)	19.82 17.30 (3, 8, 15, 27, 54)	21.23 18.56 (3, 8, 16, 28, 58)	
5	21.22 19.10 (3, 8, 15, 29, 59)	23.14 20.86 (3, 8, 17, 31, 65)	25.10 22.69 (3, 9, 18, 34, 70)	27.25 24.67 (4, 10, 20, 37, 76)		
7	29.27 27.11 (3, 10, 21, 40, 83)	32.22 29.92 (4, 11, 23, 44, 92)	35.53 33.07 (4, 12, 25, 48, 101)			
9	41.75 39.57 (4, 14, 30, 57, 121)	46.88 44.55 (5, 15, 33, 64, 136)				
11	62.38 60.18 (5, 20, 44, 86, 182)					
13						
15						
⋮						
53						

Recall that the reason why there are so many open cells is because the values of h is taken to satisfy $h \le \frac{n(n+1)}{2} - k$. For example, for $k = 11$ the reference value h is taken to be smaller than or equal to 44, since $\frac{n(n+1)}{2} - k = \frac{10(10+1)}{2} - 11 = 55 - 11 = 44$.

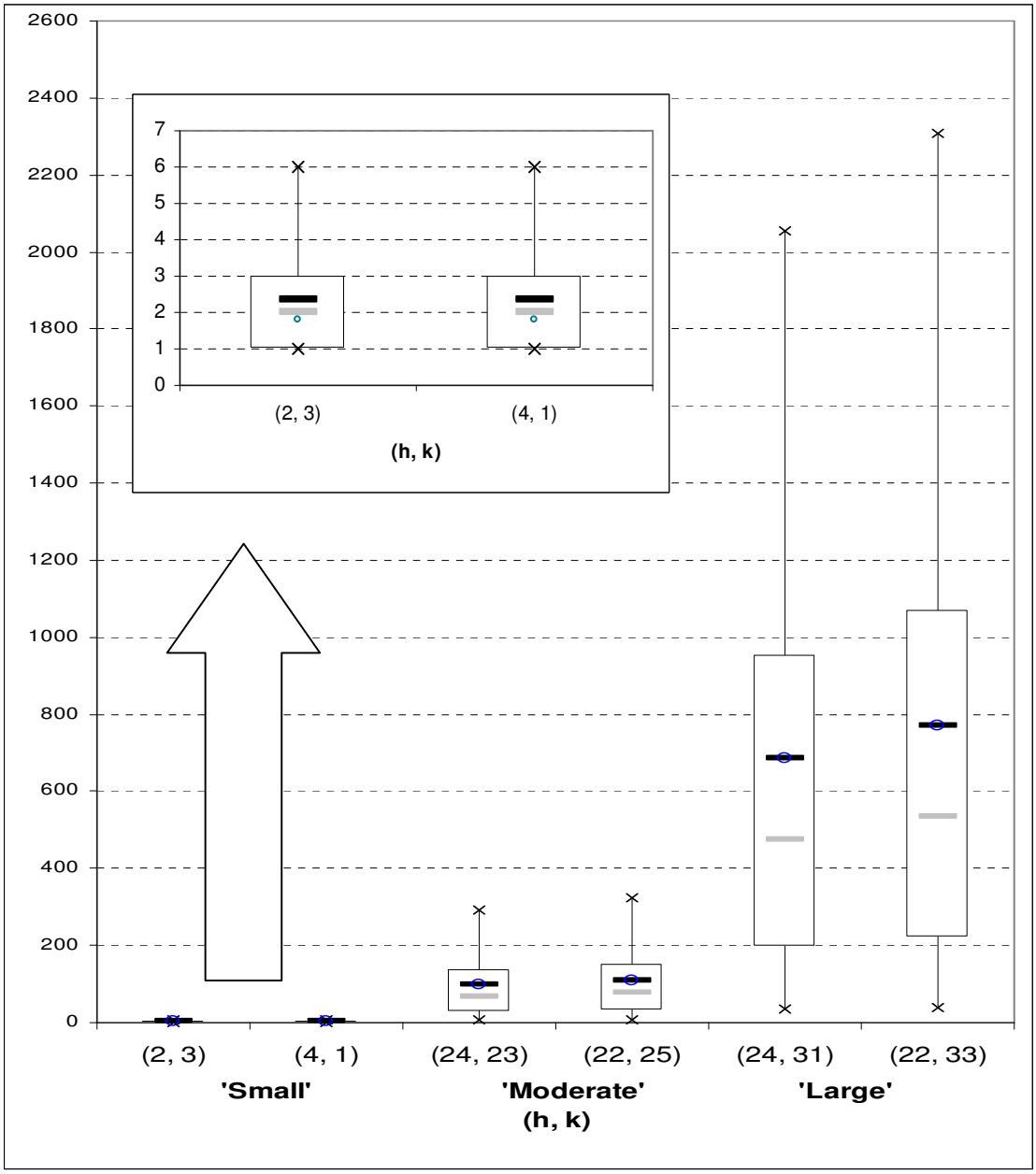

Figure 3.5. Boxplot-like graphs for the in-control run length distribution of various upper one-sided CUSUM signed-rank charts when $n = 10$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL*[*] and *MRL*, respectively[†].

[*] For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

[†] The "boxplots" are classified into 3 categories, namely small ($h+k \leq 25$), moderate ($25 < h+k \leq 50$) and large ($h+k > 50$).

Example 3.5

An upper one-sided CUSUM signed-rank chart for the Montgomery (2001) piston ring data

We conclude this sub-section by illustrating the upper one-sided CUSUM signed-rank chart using the Montgomery (2001) piston ring data. Recall that the dataset contains 15 samples (each of size 5). For illustration take $k = 3$ and $h = 8$. From Table 3.16 it can be seen that the in-control average run length equals 8.13 when $(h,k) = (8,3)$. Generally, one chooses the chart constants so that a specified in-control average run length, such as 500, or 370, is obtained. Taking this into consideration, an in-control average run length of 8.13 is considered small. Recall that unless the sample size n is 10 or more, the signed-rank chart is somewhat unattractive (from an operational point of view) in SPC applications. The plotting statistics for the Shewhart signed-rank chart (SR_i for $i = 1,2,...,15$) are given in the second row of Table 3.19. The upper one-sided CUSUM plotting statistics (S_i^+ for $i = 1,2,...,15$) are given in the last row of Table 3.19.

Table 3.19. SR_i and S_i^+ values for the piston ring data in Montgomery (2001)*.

Sample No:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$\boldsymbol{SR_i}$	8	4	-14	7	-3	9	10	-6	12	14	4	15	15	15	14
$\boldsymbol{S_i^+}$	5	6	0	4	0	6	13	4	13	24	25	37	49	61	72

To illustrate the calculations, consider sample number 1. The equation for the plotting statistic is $S_1^+ = \max[0, S_0^+ + SR_1 - k] = \max[0,0+8-3] = \max[0,5] = 5$ where a signaling event occurs for the first i such that $S_i^+ \geq h$, that is, $S_i^+ \geq 8$. The graphical display of the upper one-sided CUSUM signed-rank chart is shown in Figure 3.6.

* The values in Table 3.19 were generated using Microsoft Excel.

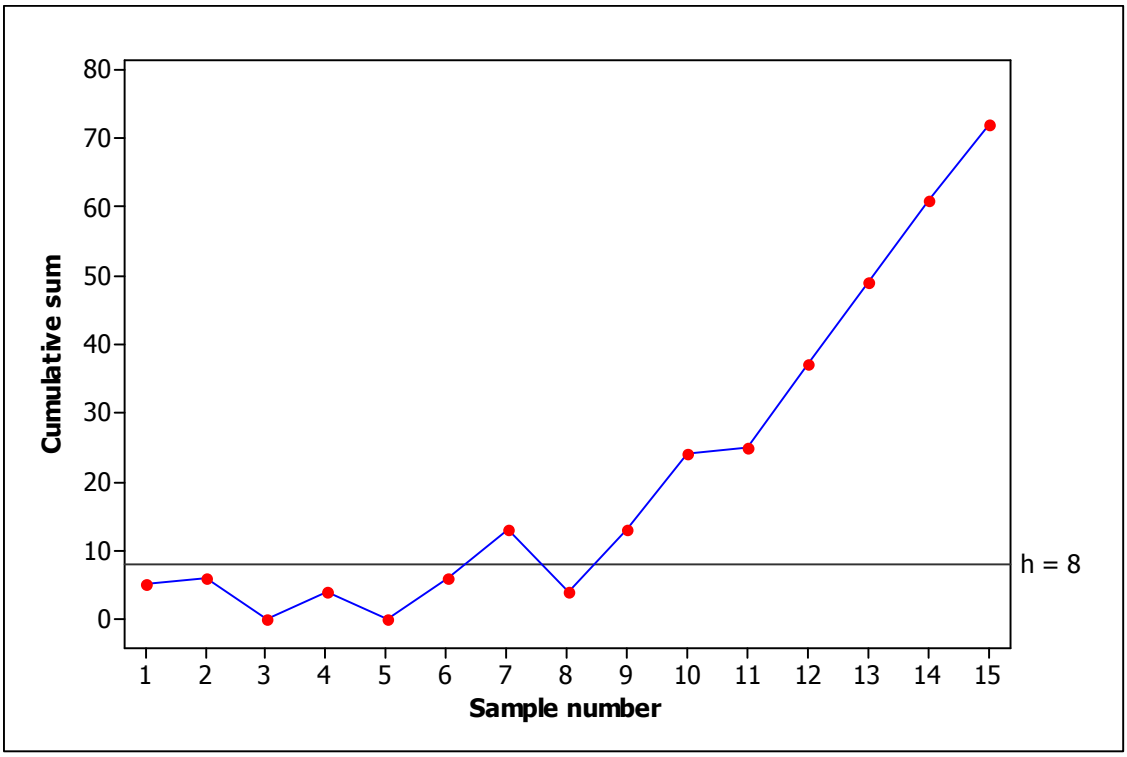

Figure 3.6. The upper one-sided CUSUM signed-rank chart for the Montgomery (2001) piston ring data.

The upper one-sided CUSUM signed-rank chart signals at sample 7, indicating a most likely positive deviation from the known target value θ_0. The action taken following an out-of-control signal on a CUSUM chart is identical to that with any control chart. A search for assignable causes should be done, corrective action should be taken (if required) and, following this, the CUSUM is reset to zero.

3.3.2.2. Lower one-sided control charts

The time that the procedure signals is the first time such that the finite-state Markov chain S_t^- enters the state ς_0 where the state space is given by $\Omega^- = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$ with $-h = \varsigma_0 < ... < \varsigma_{r+s-1} = 0$, $S_0^- = 0$ and

$$S_t^- = \max\left\{-h, \min\left\{0, S_{t-1}^- + SR_t + k\right\}\right\}. \tag{3.4}$$

Example 3.6

A lower one-sided CUSUM signed-rank chart where the sample size is even (*n*=4)

The statistical properties of a lower one-sided CUSUM signed-rank chart with a decision interval of 6 ($h = 6$), a reference value of 2 ($k = 2$) and a sample size of 4 ($n = 4$) is examined. For n even, the reference value is taken to be even, because this leads to the sum $\sum(SR_i - k)$ being equal to even values which reduces the size of the state space for the Markov chain. For $h = 6$ we have $\Omega^- = -h(2)0 = \{-6,-4,-2,0\}$. The state space is calculated using equation (3.4) and the calculations are shown in Table 3.20.

Table 3.20. Calculation of the state space when $h = 6$, $k = 2$ and $n = 4$.

SR_t	$S_{t-1}^- + SR_t + k$	$\min\{0, S_{t-1}^- + SR_t + k\}$	$S_t^- = \max\{-h, \min\{0, S_{t-1}^- + SR_t + k\}\}$
-10	-8*	-8	-6
-8	-6	-6	-6
-6	-4	-4	-4
-4	-2	-2	-2
-2	0	0	0
0	2	0	0
2	4	0	0
4	6	0	0
6	8	0	0
8	10	0	0
10	12	0	0

Table 3.21. Classification of the states.

State number	Description of the state	Absorbent (A)/ Non-absorbent (NA)
0	$S_t^+ = 0$	NA
1	$S_t^+ = -2$	NA
2	$S_t^+ = -4$	NA
3	$S_t^+ = -6$	A

* Note: Since only the state space needs to be described, S_{t-1}^- can be any value from Ω^- and we therefore take, without loss of generality, $S_{t-1}^- = 0$. Any other possible value for S_{t-1}^- would lead to the same Ω^-.

From Table 3.21 we see that there are three non-absorbent states, i.e. $r=3$, and one absorbent state, i.e. $s=1$. Therefore, the corresponding TPM will be a $(r+s)\times(r+s)=4\times4$ matrix. It can be shown (see Table 3.22) that the TPM is given by

$$TPM_{4\times4}=\begin{pmatrix} p_{00} & p_{0(-2)} & p_{0(-4)} & p_{0(-6)} \\ p_{(-2)0} & p_{(-2)(-2)} & p_{(-2)(-4)} & p_{(-2)(-6)} \\ p_{(-4)0} & p_{(-4)(-2)} & p_{(-4)(-4)} & p_{(-4)(-6)} \\ p_{(-6)0} & p_{(-6)(-2)} & p_{(-6)(-4)} & p_{(-6)(-6)} \end{pmatrix}=\left(\begin{array}{cccc|c} 11/16 & 2/16 & 1/16 & | & 2/16 \\ 9/16 & 2/16 & 2/16 & | & 3/16 \\ 7/16 & 2/16 & 2/16 & | & 5/16 \\ - & - & - & - & - \\ 0 & 0 & 0 & | & 1 \end{array}\right)=\left(\begin{array}{ccc} Q_{3\times3} & | & \underline{p}_{3\times1} \\ - & - & - \\ \underline{0}'_{1\times3} & | & 1_{1\times1} \end{array}\right)$$

where the essential transition probability sub-matrix $Q_{3\times3}:(NA\rightarrow NA)$ is an $r\times r=3\times3$ matrix, $\underline{p}_{3\times1}:(NA\rightarrow A)$ is an $(r+s-1)\times1=3\times1$ column vector, $\underline{0}'_{1\times3}:(A\rightarrow NA)$ is a $1\times(r+s-1)=1\times3$ row vector and $1_{1\times1}:(A\rightarrow A)$ represents the scalar value one.

The one-step transition probabilities are calculated by substituting SR_t in expression (3.4) by $2T^{+}-\dfrac{n(n+1)}{2}$ and substituting values for h, k, S_t^{-} and S_{t-1}^{-}. The calculation of the one-step transition probabilities are given for illustration in Table 3.22.

Recall that the probabilities in the last column of the TPM are calculated using the fact that $\sum_{j\in\Omega}p_{ij}=1\quad\forall i$ (see equation (2.18)). Therefore,

$p_{0(-6)}=1-(p_{00}+p_{0(-2)}+p_{0(-4)})=1-(11/16+2/16+1/16)=2/16$;

$p_{(-2)(-6)}=1-(p_{(-2)0}+p_{(-2)(-2)}+p_{(-2)(-4)})=1-(9/16+2/16+2/16)=3/16$;

$p_{(-4)(-6)}=1-(p_{(-4)0}+p_{(-4)(-2)}+p_{(-4)(-4)})=1-(7/16+2/16+2/16)=5/16$;

$p_{(-6)(-6)}=1-(p_{(-6)0}+p_{(-6)2}+p_{(-6)4})=1-(0+0+0)=1$.

Table 3.22. The calculation of the transition probabilities when $h=6$, $k=2$ and $n=4$.

p_{00} $= P(S_t = 0 \mid S_{t-1} = 0)$ $= P(\max\{-6, \min\{0,0+SR_t+2\}\} = 0)$ $= P(\min\{0,0+SR_t+2\} = 0)$ $= P(SR_t + 2 \geq 0)$ $= P(SR_t \geq -2)$ $= P(2T^+ - 10 \geq -2)$ $= P(T^+ \geq 4)$ $= 1 - P(T^+ \leq 3)$ $= \frac{11}{16}$	$p_{0(-2)}$ $= P(S_t = -2 \mid S_{t-1} = 0)$ $= P(\max\{-6, \min\{0,0+SR_t+2\}\} = -2)$ $= P(\min\{0, SR_t+2\} = -2)$ $= P(SR_t + 2 = -2)$ $= P(SR_t = -4)$ $= P(2T^+ - 10 = -4)$ $= P(T^+ = 3)$ $= \frac{2}{16}$	$p_{0(-4)}$ $= P(S_t = -4 \mid S_{t-1} = 0)$ $= P(\max\{-6, \min\{0,0+SR_t+2\}\} = -4)$ $= P(\min\{0, SR_t+2\} = -4)$ $= P(SR_t + 2 = -4)$ $= P(SR_t = -6)$ $= P(2T^+ - 10 = -6)$ $= P(T^+ = 2)$ $= \frac{1}{16}$
$p_{(-2)0}$ $= P(S_t = 0 \mid S_{t-1} = -2)$ $= P(\max\{-6, \min\{0,-2+SR_t+2\}\} = 0)$ $= P(\min\{0,-2+SR_t+2\} = 0)$ $= P(SR_t \geq 0)$ $== P(2T^+ - 10 \geq 0)$ $= P(T^+ \geq 5)$ $= 1 - P(T^+ \leq 4)$ $= \frac{9}{16}$	$p_{(-2)(-2)}$ $= P(S_t = -2 \mid S_{t-1} = -2)$ $= P(\max\{-6, \min\{0,-2+SR_t+2\}\} = -2)$ $= P(\min\{0, SR_t\} = -2)$ $= P(SR_t = -2)$ $== P(2T^+ - 10 = -2)$ $= P(T^+ = 4)$ $= \frac{2}{16}$	$p_{(-2)(-4)}$ $= P(S_t = -4 \mid S_{t-1} = -2)$ $= P(\max\{-6, \min\{0,-2+SR_t+2\}\} = -4)$ $= P(\min\{0, SR_t\} = -4)$ $= P(SR_t = -4)$ $= P(2T^+ - 10 = -4)$ $= P(T^+ = 3)$ $= \frac{2}{16}$
$p_{(-4)0}$ $= P(S_t = 0 \mid S_{t-1} = -4)$ $= P(\max\{-6, \min\{0,-4+SR_t+2\}\} = 0)$ $= P(\min\{0,-4+SR_t+2\} = 0)$ $= P(SR_t \geq 2)$ $== P(2T^+ - 10 \geq 2)$ $= P(T^+ \geq 6)$ $= 1 - P(T^+ \leq 5)$ $= \frac{7}{16}$	$p_{(-4)(-2)}$ $= P(S_t = -2 \mid S_{t-1} = -4)$ $= P(\max\{-6, \min\{0,-4+SR_t+2\}\} = -2)$ $= P(\min\{0, SR_t-2\} = -2)$ $= P(SR_t - 2 = -2)$ $== P(2T^+ - 10 = 0)$ $= P(T^+ = 5)$ $= \frac{2}{16}$	$p_{(-4)(-4)}$ $= P(S_t = -4 \mid S_{t-1} = -4)$ $= P(\max\{-6, \min\{0,-4+SR_t+2\}\} = -4)$ $= P(\min\{0, SR_t-2\} = -4)$ $= P(SR_t - 2 = -4)$ $= P(SR_t = -2)$ $= P(2T^+ - 10 = -2)$ $= P(T^+ = 4)$ $= \frac{2}{16}$
$p_{(-6)0}$ $= P(S_t = 0 \mid S_{t-1} = -6)$ $= 0$ *	$p_{(-6)(-2)}$ $= P(S_t = -2 \mid S_{t-1} = -6)$ $= 0$	$p_{(-6)(-4)}$ $= P(S_t = -4 \mid S_{t-1} = -6)$ $= 0$

* The probability equals zero, because it is impossible to go from an absorbent state to a non-absorbent state.

The formulas of the moments and some characteristics of the run length distribution have been studied by Fu, Spiring and Xie (2002) and Fu and Lou (2003) – see equations (2.41) to (2.45). By substituting $\underline{\xi}_{1\times 3} = (1 \quad 0 \quad 0)$, $Q_{3\times 3} = \begin{pmatrix} 11/16 & 2/16 & 1/16 \\ 9/16 & 2/16 & 2/16 \\ 7/16 & 2/16 & 2/16 \end{pmatrix}$ and $\underline{1}_{3\times 1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ into these equations, we obtain the following:

$$ARL = E(N) = \underline{\xi}(I-Q)^{-1}\underline{1} = 6.81$$

$$E(N^2) = \underline{\xi}(I+Q)(I-Q)^{-2}\underline{1} = 83.64$$

$$SDRL = \sqrt{Var(N)} = \sqrt{E(N^2) - (E(N))^2} = 6.11$$

5^{th} percentile $= \rho_5 = 1$

25^{th} percentile $= \rho_{25} = 2$

Median $= 50^{th}$ percentile $= \rho_{50} = 5$

75^{th} percentile $= \rho_{75} = 9$

95^{th} percentile $= \rho_{95} = 19$

The in-control average run length (ARL_0^-) values, standard deviation of the run length ($SDRL$) values and percentiles for the lower one-sided CUSUM signed-rank chart are exactly the same as for the upper one-sided CUSUM signed-rank chart, since the one-step transition probabilities matrices are the same. Therefore, the in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile values for the upper one-sided CUSUM signed-rank chart will also hold for the lower one-sided CUSUM signed-rank chart.

Example 3.7

A lower one-sided CUSUM signed-rank chart for the Montgomery (2001) piston ring data

We conclude this sub-section by illustrating the lower one-sided CUSUM signed-rank chart using the Montgomery (2001) piston ring data. Recall that the dataset contains 15 samples (each of size 5). For illustration take $k = 3$ and $h = 8$. From Table 3.16 it can be seen that the in-control average run length equals 8.13 when $(h,k) = (8,3)$. Generally, one chooses the chart constants so that a specified in-control average run length, such as 500, or 370, is obtained. Taking this into consideration, an in-control average run length of 8.13 is considered small. Recall that unless the sample size n is 10 or more, the signed-rank chart is somewhat unattractive (from an operational point of view) in SPC applications.

The plotting statistics for the Shewhart signed-rank chart (SR_i for $i = 1,2,...,15$) are given in the second row of Table 3.23. The lower one-sided CUSUM plotting statistics (S_i^- for $i = 1,2,...,15$) are given in the last row of Table 3.23.

Table 3.23. SR_i and S_i^- values for the piston ring data in Montgomery (2001)*.

Sample No:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
$\boldsymbol{SR_i}$	8	4	-14	7	-3	9	10	-6	12	14	4	15	15	15	14
$\boldsymbol{S_i^-}$	0	0	-11	-1	-1	0	0	-3	0	0	0	0	0	0	0

To illustrate the calculations, consider sample number 1. The equation for the plotting statistic S_1^- is

$$S_1^{-*} = \max[0, S_0^{-*} - SR_1 - k] = \max[0,0-8-3] = \max[0,-11] = 0 \qquad (3.5)$$

or

$$S_1^- = \min[0, S_0^- + SR_1 + k] = \min[0,0+8+3] = \min[0,11] = 0 \qquad (3.6)$$

* The values in Table 3.23 were generated using Microsoft Excel.

A signaling event occurs for the first i such that $S_i^{-*} \geq h$, that is, $S_i^{-*} \geq 8$ if expression (3.5) is used or $S_i^{-} \leq -h$, that is, $S_i^{-} \leq -8$ if expression (3.6) is used. The graphical display of the lower one-sided CUSUM signed-rank chart is shown in Figure 3.7.

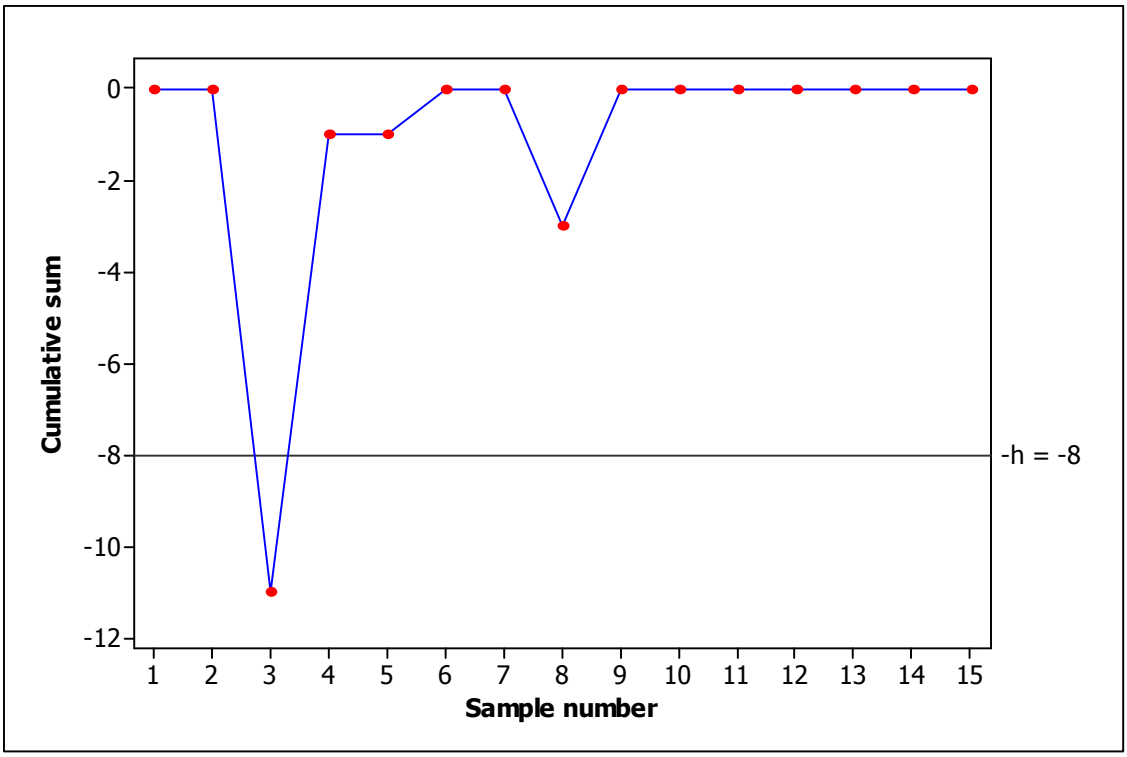

Figure 3.7. The lower one-sided CUSUM signed-rank chart for the Montgomery (2001) piston-ring data.

The lower one-sided CUSUM signed-rank chart signals at sample 3. Recall that the lower one-sided CUSUM sign chart did not signal at all. This emphasizes the fact that the signed-rank test is more powerful than the sign test. The question arises: Why not always use the signed-rank test if it is more powerful than the sign test? The sign test is applicable for all continuous distributions, while the assumption of symmetry must be made, in addition, for the signed-rank test. Also, the sign test applies to all percentiles while the signed-rank test is proposed only for the median.

3.3.3. Two-sided control charts

Recall that for the upper one-sided CUSUM signed-rank chart we use

$$S_t^+ = \min\left\{h, \max\{0, SR_t - k + S_{t-1}^+\}\right\} \text{ for } t = 1,2,... \tag{3.7}$$

For a lower one-sided CUSUM signed-rank chart we use

$$S_t^- = \max\left\{-h, \min\{0, SR_t + k + S_{t-1}^-\}\right\} \text{ for } t = 1,2,... \tag{3.8}$$

For the two-sided scheme the two one-sided schemes are performed simultaneously. The corresponding two-sided CUSUM chart signals for the first n at which either one of the two inequalities is satisfied, that is, either $S_t^+ \geq h$ or $S_t^- \leq -h$. Starting values are typically chosen to equal zero, that is, $S_0^+ = S_0^- = 0$. The two-sided scheme signals at N where

$$N = \min_t\left\{t : S_t^+ \geq h \quad \text{or} \quad S_t^- \leq -h\right\} \tag{3.9}$$

where h is a positive integer.

The two-sided CUSUM scheme can be represented by a Markov chain with states corresponding to the possible combinations of values of S_t^+ and S_t^-. The states corresponding to values for which a signal is given by the CUSUM scheme are called absorbing states. Clearly, there are two absorbing states ($s = 2$) since the chart signals when S_t^+ falls on or above h or when S_t^- falls on or below $-h$. The probability of going from an absorbing state to the same absorbing state is equal to one, because once an absorbing state is entered, it is never left. The transient states are the remaining states for which eventual return is uncertain. Let r denote the number of remaining states, i.e. r denotes the number of transient (non-absorbing) states. Clearly, in total there are $r+s$ states and therefore the corresponding TPM will be an $(r+s)\times(r+s)$ matrix.

The time that the procedure signals is the first time such that the finite-state Markov chain enters the state ς_0 or ς_{r+s-1} where the state space is given by $\Omega = \Omega^+ \cup \Omega^- = \{\varsigma_0, \varsigma_1, ..., \varsigma_{r+s-1}\}$ with $-h = \varsigma_0 < ... < \varsigma_{r+s-1} = h$.

Example 3.8

A two-sided CUSUM signed-rank chart where the sample size is even (*n*=4)

The statistical properties of a two-sided CUSUM signed-rank chart with a decision interval of 4 ($h = 4$), a reference value of 2 ($k = 2$) and a sample size of 4 ($n = 4$) is examined. Let Ω denote the state space for the two-sided chart. Ω is the union of the state space for the upper one-sided chart $\Omega^+ = \{0,2,4\}$ and the state space for the lower one-sided chart $\Omega^- = \{-4,-2,0\}$. Therefore, $\Omega = \Omega^+ \cup \Omega^- = \{-4,-2,0\} \cup \{0,2,4\} = \{-4,-2,0,2,4\} = \{\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4\}$ with $-h = \varsigma_0 < \varsigma_1 < \varsigma_2 < \varsigma_3 < \varsigma_4 = h$.

Table 3.24. Classification of the states.

State number	Values of the CUSUM statistic(s)	Absorbing (A) Non-absorbing (NA)
0	$S_t^- = 0$ and $S_t^+ = 0$	NA
1	$S_t^- = 2$ or $S_t^+ = 2$ *	NA
2	$S_t^- = -2$ or $S_t^+ = -2$ †	NA
3	$S_t^- = 4$ or $S_t^+ = 4$ ‡	A
4	$S_t^- = -4$ or $S_t^+ = -4$ §	A

From Table 3.24 we see that there are three non-absorbing states, i.e. $r = 3$, and two absorbing states, i.e. $s = 2$. Therefore the corresponding TPM will be a (5×5) matrix. The layout of the TPM is as follows. There are three transient states and two absorbing states. By

* Moving from state 0 to state 1 can happen when either the upper cumulative sum or the lower cumulative sum equals 2. But the lower cumulative sum can not equal 2 since by definition the lower cumulative sum can only take on integer values smaller than or equal to zero. Therefore, we only use the probability that the upper cumulative sum equals 2 in the calculation of the probabilities in the TPM.

† Moving from state 0 to state 2 can happen when either the upper cumulative sum or the lower cumulative sum equals -2. But the upper cumulative sum can not equal -2 since by definition the upper cumulative sum can only take on integer values greater than or equal to zero. Therefore, we only use the probability that the lower cumulative sum equals -2 in the calculation of the probabilities in the TPM.

‡ A similar argument to the argument in the first footnote on this page holds. Therefore, we only use the probability that the upper cumulative sum equals 4 in the calculation of the probabilities in the TPM.

§ A similar argument to the argument in the second footnote on this page holds. Therefore, we only use the probability that the lower cumulative sum equals -4 in the calculation of the probabilities in the TPM.

convention we first list the non-absorbing states and then we list the absorbing states. In column one we compute the probability of moving from state i to state 0, for all i. Note that the process reaches state 0 when both the upper and the lower cumulative sums equal zero. In columns two and three, we compute the probabilities of moving from state i to the remaining non-absorbing states, for all i. Finally, in the remaining two columns we compute the probabilities of moving from state i to the absorbing states, for all i. Thus, the TPM can be conveniently partitioned into 4 sections given by

$$TPM_{5\times5} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \left(\begin{array}{cccccc} 6/16 & 2/16 & 2/16 & | & 3/16 & 3/16 \\ 4/16 & 2/16 & 2/16 & | & 5/16 & 3/16 \\ 4/16 & 2/16 & 2/16 & | & 3/16 & 5/16 \\ - & - & - & - & - & - \\ 0 & 0 & 0 & | & 1 & 0 \\ 0 & 0 & 0 & | & 0 & 1 \end{array}\right) = \begin{pmatrix} Q_{3\times3} & | & C_{3\times2} \\ - & - & - \\ Z_{2\times3} & | & I_{2\times2} \end{pmatrix}$$

where $Q_{3\times3} : (NA \to NA)$ is an $r \times r = 3\times3$ matrix, $C_{3\times2} : (NA \to A)$ is an $r \times s = 3\times2$ matrix, $Z_{2\times3} : (A \to NA)$ is an $s \times r = 2\times3$ matrix and $I_{2\times2} : (A \to A)$ is an $s \times s = 2\times2$ matrix .

The calculation of the elements of the TPM is illustrated next. Note that this essentially involves the calculation of the matrices Q and C. First consider the transient states. Note that the process moves to state 0, i.e., $j = 0$, when both the upper *and* the lower cumulative sums equal 0. Thus the required probability of moving to 0, from any other transient state, is the probability of an *intersection* of two sets involving values of the upper and the lower CUSUM statistics, respectively. On the other hand, the probability of moving to any state $j \neq 0$, from any other state, is the probability of a *union* of two sets involving values of the upper and the lower CUSUM statistics, respectively. However, one of these two sets is empty so that the required probability is the probability of only the non-empty set.

The calculation of the entry in the first row and the first column of the TPM, p_{00}, will be discussed in detail. This is the probability of moving from state 0 to state 0 in one step at time *t*. As we just described, this can happen only when the upper *and* the lower cumulative sums both equal 0 at time t. For the upper one-sided CUSUM p_{00} is the probability that the upper CUSUM

equals 0 at time t, given that the upper CUSUM equaled 0 at time $t-1$, that is, $P(S_t^+ = 0 \mid S_{t-1}^+ = 0)$. For the lower one-sided procedure p_{00} is the probability that the lower CUSUM equals 0 at time t, given that the lower CUSUM equaled 0 at time $t-1$, that is, $P(S_t^- = 0 \mid S_{t-1}^- = 0)$. For the two-sided procedure the two one-sided procedures are performed simultaneously. Therefore we have that $p_{00} = P\left(\left\{S_t^+ = 0 \mid S_{t-1}^+ = 0\right\} \cap \left\{S_t^- = 0 \mid S_{t-1}^- = 0\right\}\right)$. We have that

$$
\begin{aligned}
& p_{00} \\
& = P\left(\left\{S_t^+ = 0 \mid S_{t-1}^+ = 0\right\} \cap \left\{S_t^- = 0 \mid S_{t-1}^- = 0\right\}\right)
\end{aligned}
$$

this is computed by substituting in values for h, k, S_t^+, S_{t-1}^+, S_t^- and S_{t-1}^- into (3.7) and (3.8)

$$
\begin{aligned}
& = P\left(\left\{\min\left\{4, \max\left\{0, SR_t - 2 + 0\right\}\right\} = 0\right\} \cap \left\{\max\left\{-4, \min\left\{0, SR_t + 2 + 0\right\}\right\} = 0\right\}\right) \\
& = P\left(\left\{\max\left\{0, SR_t - 2 + 0\right\} = 0\right\} \cap \left\{\min\left\{0, SR_t + 2 + 0\right\} = 0\right\}\right) \\
& = P\left(\left(SR_t - 2 \le 0\right) \cap \left(SR_t + 2 + 0 \ge 0\right)\right) \\
& = P\left(\left(SR_t \le 2\right) \cap \left(SR_t \ge -2\right)\right)
\end{aligned}
$$

recall that $SR_t = 2T^+ - \dfrac{n(n+1)}{2}$ where T^+ is the Wilcoxon signed-rank statistic

$$
\begin{aligned}
& = P\left(\left(2T^+ - 10 \le 2\right) \cap \left(2T^+ - 10 \ge -2\right)\right) \\
& = P\left(\left(T^+ \le 6\right) \cap \left(T^+ \ge 4\right)\right) \\
& = P(T^+ = 4) + P(T^+ = 5) + P(T^+ = 6) \\
& = {}^{6}\!/_{16}.
\end{aligned}
$$

The remaining entries of the TPM can be calculated similarly. In doing so, we find that

$$
TPM = \begin{pmatrix}
{}^{6}\!/_{16} & {}^{2}\!/_{16} & {}^{2}\!/_{16} & {}^{3}\!/_{16} & {}^{3}\!/_{16} \\
{}^{4}\!/_{16} & {}^{2}\!/_{16} & {}^{2}\!/_{16} & {}^{5}\!/_{16} & {}^{3}\!/_{16} \\
{}^{4}\!/_{16} & {}^{2}\!/_{16} & {}^{2}\!/_{16} & {}^{3}\!/_{16} & {}^{5}\!/_{16} \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}.
$$

Using the TPM the *ARL* can be calculated using $ARL = \underline{\xi}(I - Q)^{-1}\underline{1}$. A well-known concern is that important information about the performance of a control chart can be missed when only examining the *ARL* (this is especially true when the process distribution is skewed). Various authors, see for example, Radson and Boyd (2005) and Chakraborti (2007), have

suggested that one should examine a number of percentiles, including the median, to get the complete information about the performance of a control chart. Therefore, we now also consider percentiles. The calculation of these percentiles is shown in Table 3.25 for illustration purposes. The first column of Table 3.25 contains the values that the run length variable (N) can take on.

Table 3.25. Calculation of the percentiles when $h = 4$, $k = 2$ and $n = 4$ [*].

N	$P(N \le l)$	The 5th, 25th, 50th, 75th and 95th percentiles
1	0.3750000	$\rho_{0.25} = 1$ (smallest integer such that cdf is at least 0.05 and 0.25)
2	0.6406250	$\rho_{0.5} = 2$ (smallest integer such that cdf is at least 0.50)
3	0.7949219	$\rho_{0.75} = 3$ (smallest integer such that cdf is at least 0.75)
4	0.8830566	
5	0.9333191	
6	0.9619789	$\rho_{0.95} = 6$ (smallest integer such that cdf is at least 0.95)
7[†]	0.9783206	

The formulas of the moments and some characteristics of the run length distribution have been studied by Fu, Spiring and Xie (2002) and Fu and Lou (2003) – see equations (2.41) to (2.45). By substituting $\underline{\xi}_{1\times 3} = (1 \quad 0 \quad 0)$, $Q_{3\times 3} = \frac{1}{16}\begin{pmatrix} 6 & 2 & 2 \\ 4 & 2 & 2 \\ 4 & 2 & 2 \end{pmatrix}$ and $\underline{1}_{3\times 1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ into these equations, we obtain the following:

$$ARL = E(N) = \underline{\xi}(I - Q)^{-1}\underline{1} = 2.46$$

$$E(N^2) = \underline{\xi}(I + Q)(I - Q)^{-2}\underline{1} = 9.28$$

$$SDRL = \sqrt{Var(N)} = \sqrt{E(N^2) - (E(N))^2} = 1.79$$

5^{th} percentile $= p_{0.05} = 1$

25^{th} percentile $= p_{0.25} = 1$

Median $= 50^{th}$ percentile $= p_{0.5} = 2$

[*] See SAS Program 7 in Appendix B for the calculation of the values in Table 3.25.
[†] The value of the run length variable is only shown up to $N = 7$ for illustration purposes.

75^{th} percentile $= p_{0.75} = 3$

95^{th} percentile $= p_{0.95} = 6$

Other values of h, k and n were also considered and the results are given in Table 3.26.

Table 3.26. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile values* for samples of size $n = 4$ and various values of h and k for the two-sided CUSUM signed-rank chart†.

k	h				
	2	**4**	**6**	**8**	**10**
0	1.14 0.40 (1, 1, 1, 1, 2)	1.52 0.81 (1, 1, 1, 2, 3)	2.14 1.51 (1, 1, 1, 3, 5)	2.74 2.19 (1, 1, 2, 3, 7)	3.62 2.80 (1, 1, 2, 4, 10)
2	1.60 0.98 (1, 1, 1, 2, 4)	2.46 1.79 (1, 1, 2, 3, 6)	3.41 2.66 (1, 1, 2, 4, 9)	5.09 4.07 (1, 2, 4, 7, 13)	
4	2.67 2.11 (1, 1, 2, 3, 7)	3.87 3.29 (1, 1, 3, 5, 10)	6.64 5.92 (1, 2, 5, 9, 18)		
6	4.00 3.46 (1, 1, 3, 5, 11)	7.53 6.95 (1, 3, 5, 10, 21)			
8	8.00 7.45 (1, 2, 5, 11, 23)				

Values of k and h are restricted to be integers so that the Markov chain approach could be employed to obtain exact values for the average run length. In order to allow for the possibility of stopping after one group, the values of h is taken to satisfy $h \le \frac{n(n+1)}{2} - k$. For example, for $n = 4$ and $k = 0$, the reference value h is taken to be smaller than or equal to 10, since $\frac{n(n+1)}{2} - k = \frac{4(4+1)}{2} - 0 = 10$. The five percentiles are displayed in boxplot-like graphs in Figure 3.8 for all the (h,k)-combinations that are shaded in Table 3.26. The "boxplots" are

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.
† See SAS Program 7 in Appendix B for the calculation of the values in Table 3.26.

classified into 3 categories, namely small ($h+k \leq 4$), moderate ($5 \leq h+k \leq 8$) and large ($h+k \geq 9$).

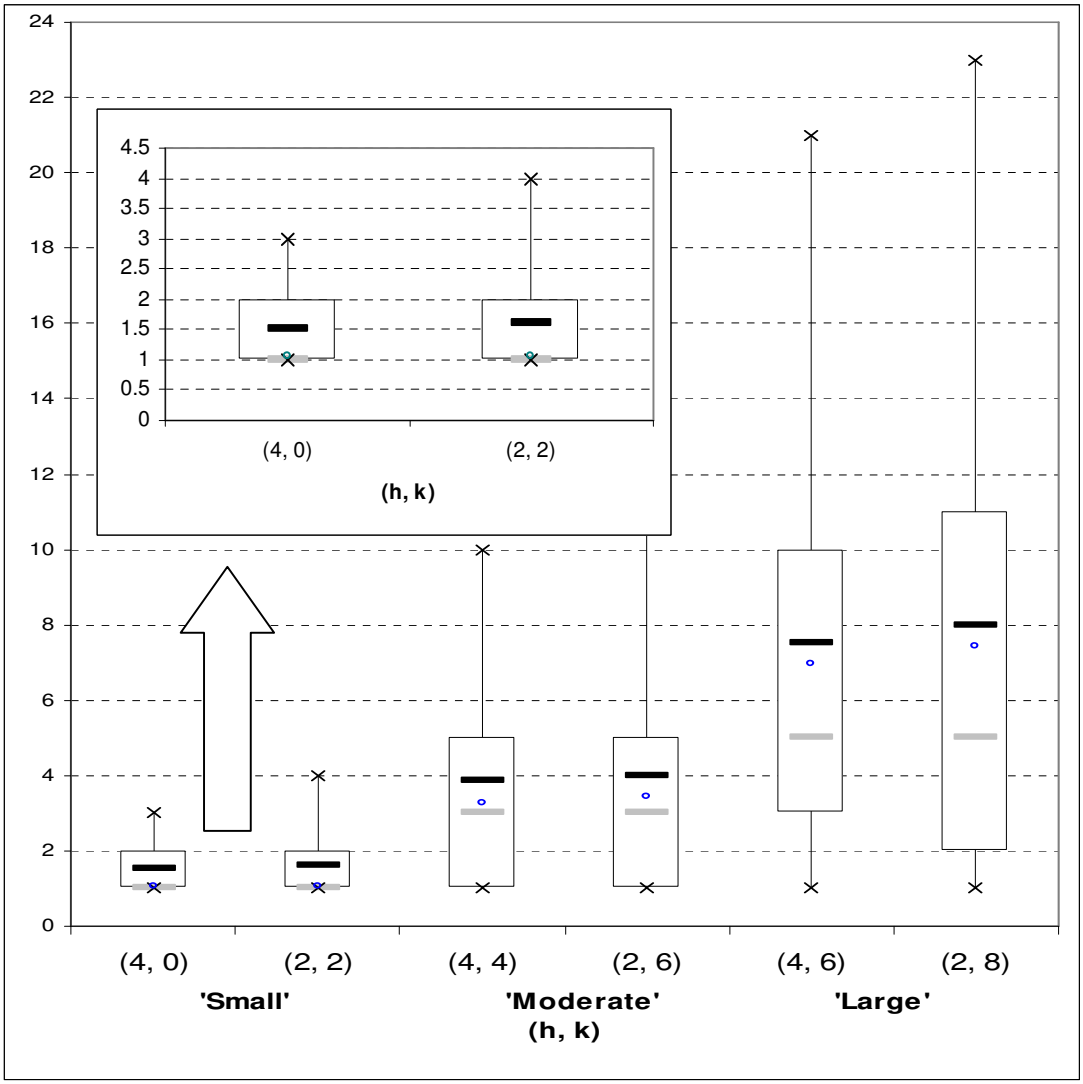

Figure 3.8. Boxplot-like graphs for the in-control run length distribution of various two-sided CUSUM signed-rank charts when $n = 4$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL*[*] and *MRL*, respectively.

[*] For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Table 3.27. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), $5^{th}, 25^{th}, 50^{th}, 75^{th}$ and 95^{th} percentile values* for samples of size $n = 5$ and various values of h and k for the two-sided CUSUM signed-rank chart†.

k	h						
	2	**4**	**6**	**8**	**10**	**12**	**14**
1	1.23 0.53 (1, 1, 1, 1, 2)	1.56 0.88 (1, 1, 1, 2, 3)	2.04 1.30 (1, 1, 2, 3, 5)	2.57 1.87 (1, 1, 2, 3, 7)	3.36 2.97 (1, 1, 3, 5, 11)	4.15 3.46 (1, 1, 3, 5, 11)	5.23 4.38 (1, 2, 4, 7, 14)
3	1.60 0.98 (1, 1, 1, 2, 4)	2.20 1.57 (1, 1, 2, 3, 5)	2.90 2.22 (1, 1, 2, 4, 7)	4.07 3.23 (1, 2, 3, 5, 10)	5.34 4.36 (1, 2, 4, 7, 14)	7.39 6.61 (1, 2, 5, 10, 21)	
5	2.29 1.71 (1, 1, 2, 3, 6)	3.12 2.54 (1, 1, 2, 4, 8)	4.72 4.05 (1, 2, 3, 6, 13)	6.52 5.77 (1, 2, 5, 9, 18)	10.09 9.11 (1, 4, 7, 14, 28)		
7	3.20 2.65 (1, 1, 2, 4, 8)	5.12 4.55 (1, 2, 4, 7, 14)	7.39 6.78 (1, 3, 5, 10, 21)	12.58 11.83 (1, 4, 9, 17, 36)			
9	5.33 4.81 (1, 2, 4, 7, 15)	7.87 7.34 (1, 3, 6, 11, 23)	14.57 13.97 (1, 5, 10, 20, 42)				
11	8.00 7.48 (1, 3, 6, 11, 23)	15.52 14.98 (1, 5, 11, 21, 45)					
13	16.00 15.49 (1, 5, 11, 22, 47)						

The five percentiles are displayed in boxplot-like graphs in Figure 3.9 for all the (h,k)-combinations that are shaded in Table 3.27. The "boxplots" are classified into 3 categories, namely small ($h+k \le 5$), moderate ($6 \le h+k \le 10$) and large ($h+k \ge 11$).

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles ($\rho_5, \rho_{25}, \rho_{50}, \rho_{75}, \rho_{95}$), respectively.
† See SAS Program 7 in Appendix B for the calculation of the values in Table 3.27.

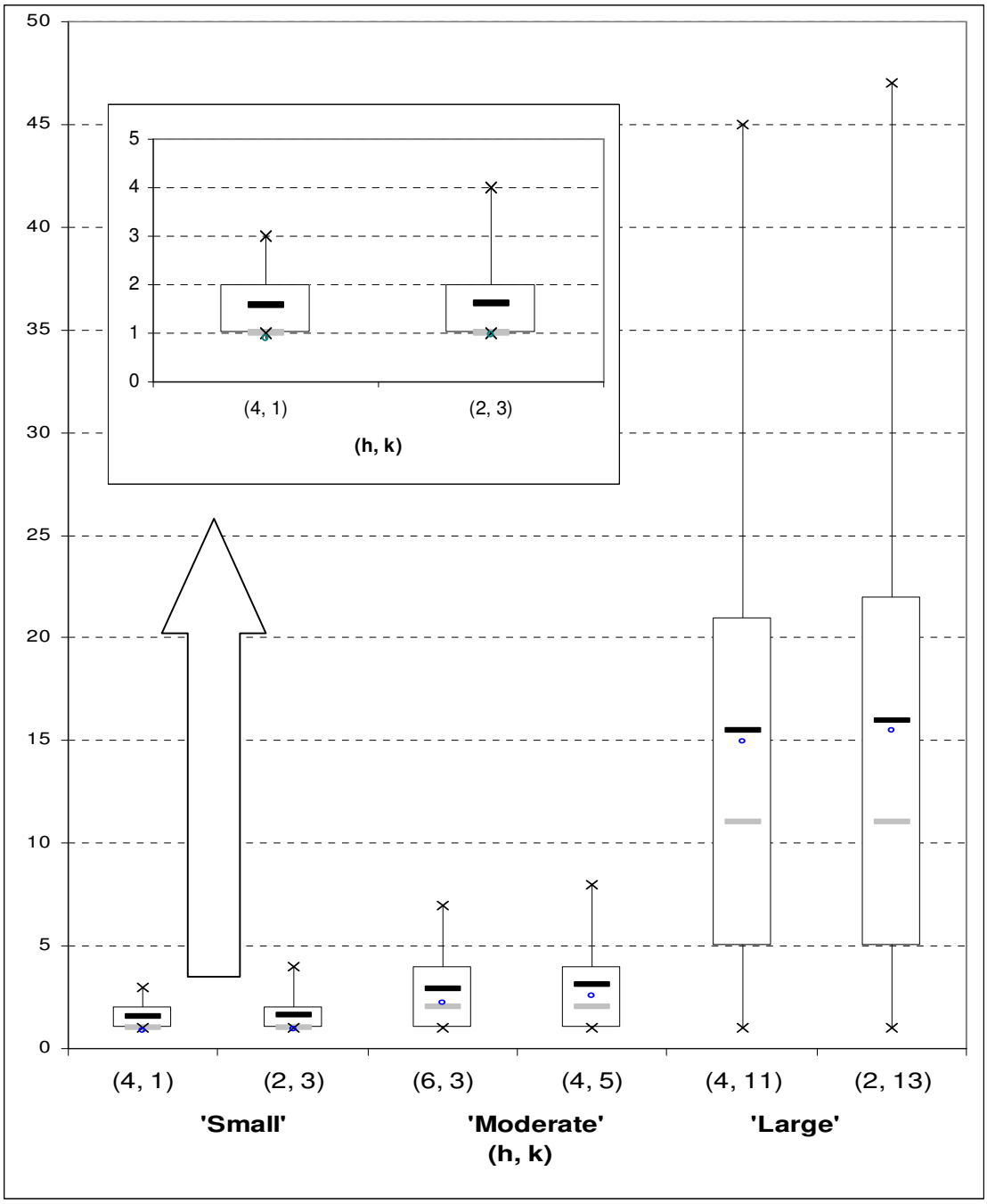

Figure 3.9. Boxplot-like graphs for the in-control run length distribution of various two-sided CUSUM signed-rank charts when $n = 5$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL*[*] and *MRL*, respectively.

[*] For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

Table 3.28. The in-control average run length (ARL_0^+), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for samples of size $n = 6$ and various values of h and k for the two-sided CUSUM signed-rank chart†.

k	h								
	2	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**
1	1.19 0.47 (1, 1, 1, 1, 2)	1.43 0.75 (1, 1, 1, 2, 3)	1.69 1.01 (1, 1, 1, 2, 4)	2.08 1.37 (1, 1, 1, 2, 5)	2.52 2.10 (1, 1, 2, 3, 6)	3.05 2.48 (1, 1, 2, 4, 8)	3.62 3.09 (1, 1, 2, 5, 10)	4.36 3.32 (1, 1, 3, 6, 11)	5.11 4.24 (1, 2, 4, 7, 14)
3	1.45 0.81 (1, 1, 1, 2, 3)	1.75 1.13 (1, 1, 1, 2, 4)	2.17 1.52 (1, 1, 2, 3, 5)	2.77 2.04 (1, 1, 2, 4, 7)	3.50 2.65 (1, 2, 3, 5, 9)	4.32 3.36 (1, 1, 3, 6, 12)	5.49 4.72 (1, 2, 4, 8, 15)	6.72 5.96 (1, 2, 5, 9, 19)	8.39 7.49 (1, 3, 6, 11, 24)
5	1.78 1.18 (1, 1, 1, 2,4)	2.25 1.65 (1, 1, 1, 2, 3, 6)	2.98 2.34 (1, 1, 2, 4, 8)	3.91 3.20 (1, 2, 3, 5, 10)	5.01 4.21 (1, 2, 4, 7, 13)	6.77 5.79 (1, 3, 5, 9, 18)	8.69 7.56 (1, 3, 6, 12, 24)	11.72 10.54 (1, 5, 11, 22, 47)	
7	2.29 1.71 (1, 1, 2, 3, 6)	3.12 2.54 (1, 1, 2, 4, 8)	4.25 3.63 (1, 2, 3, 6, 11)	5.63 4.96 (1, 2, 4, 8, 16)	8.09 7.28 (1, 3, 6, 11, 23)	10.90 9.98 (1, 4, 8, 15, 31)	16.01 14.85 (2, 5, 11, 22, 46)		
9	3.20 2.65 (1, 1, 2, 4, 8)	4.48 3.93 (1, 2, 3, 6, 12)	6.08 5.50 (1, 2, 4, 8, 17)	9.24 8.57 (1, 3, 7, 13, 26)	12.95 12.21 (1, 4, 9, 18, 37)	20.78 19.85 (2, 7, 15, 28, 60)			
11	4.57 4.04 (1, 2, 3, 6, 13)	6.32 5.78 (1, 2, 5, 9, 18)	10.02 9.44 (1, 3, 7, 14, 29)	14.44 13.82 (1, 5, 10, 20, 24)	25.13 24.38 (2, 8, 18, 35, 74)				
13	6.40 5.88 (1, 2, 5, 9, 18)	10.45 9.91 (1, 3, 7, 14, 30)	15.38 14.83 (1, 5, 11, 21, 45)	28.31 27.68 (2, 9, 20, 39, 84)					
15	10.67 10.15 (1, 3, 8, 15, 31)	15.87 15.36 (1, 5, 11, 22, 47)	30.54 29.99 (2, 9, 21, 42, 90)						
17	16.00 15.49 (1, 5, 11, 22, 47)	31.51 30.99 (2, 9, 22, 43, 93)							
19	32.00 31.50 (2, 10, 22, 44, 95)								

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.

† See SAS Program 7 in Appendix B for the calculation of the values in Table 3.28.

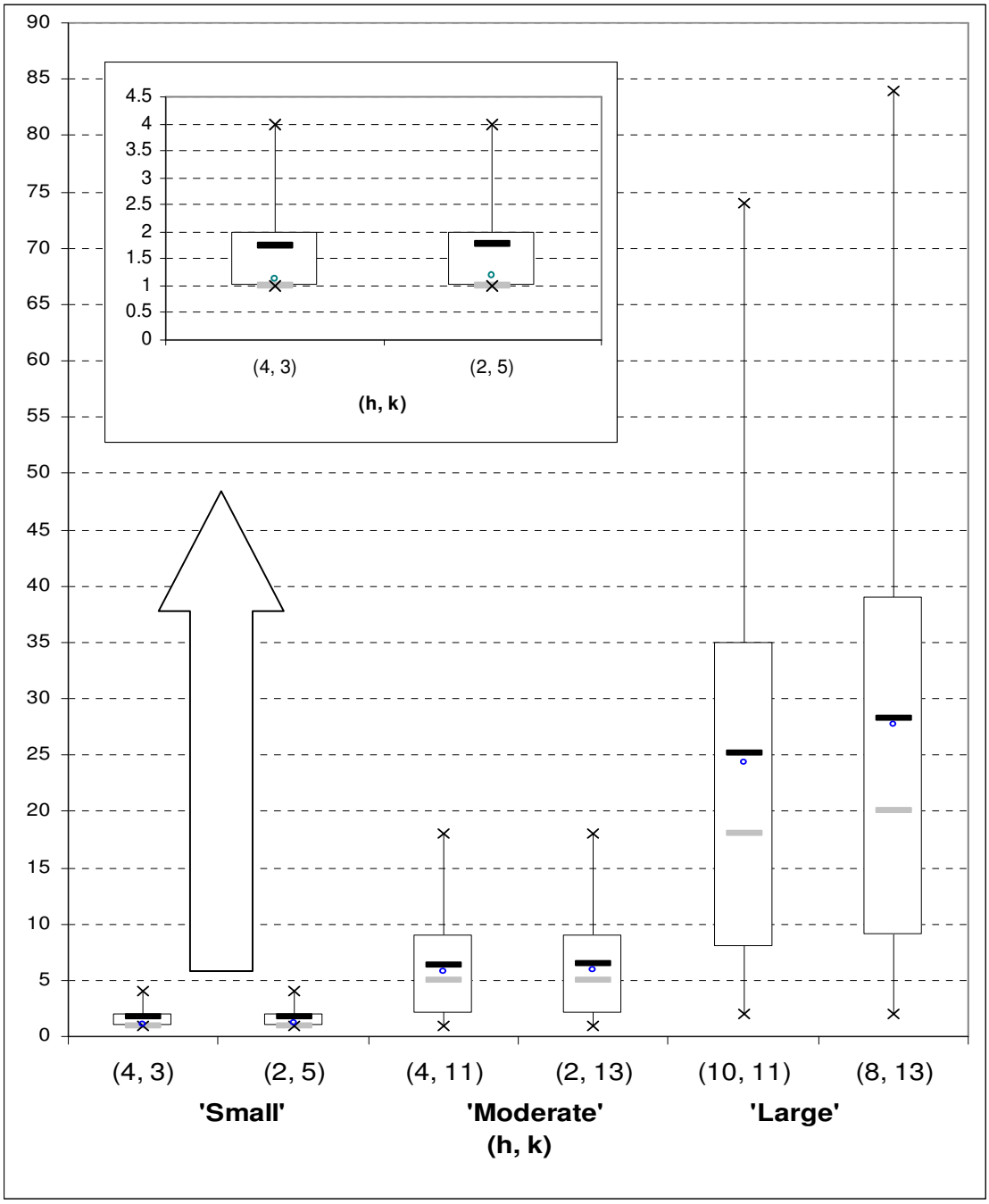

Figure 3.10. Boxplot-like graphs for the in-control run length distribution of various two-sided CUSUM signed-rank charts when $n = 6$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "○" and "▬" denote the *ARL*, *SDRL* and *MRL*, respectively*.

* The "boxplots" are classified into 3 categories, namely small ($h + k \leq 7$), moderate ($8 \leq h + k \leq 16$) and large ($h + k \geq 17$).

Table 3.29. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for samples of size $n = 10$ for $h = 2,4,...,14$ and $k = 1,3,...,23$ for the two-sided CUSUM signed-rank chart†.

k	*h*						
	2	**4**	**6**	**8**	**10**	**12**	**14**
1	1.09 0.80 (1, 1, 1, 1, 2)	1.18 0.89 (1, 1, 1, 1, 3)	1.29 1.00 (1, 1, 1, 1, 3)	1.41 1.11 (1, 1, 1, 2, 3)	1.54 1.24 (1, 1, 1, 2, 4)	1.68 1.37 (1, 1, 1, 2, 4)	1.84 1.51 (1, 1, 1, 2, 5)
3	1.18 0.90 (1, 1, 1, 1, 3)	1.30 1.01 (1, 1, 1, 1, 3)	1.42 1.14 (1, 1, 1, 2, 3)	1.56 1.27 (1, 1, 1, 2, 4)	1.72 1.42 (1, 1, 1, 2, 4)	1.90 1.59 (1, 1, 1, 2, 5)	2.09 1.76 (1, 1, 1, 3, 5)
5	1.30 1.02 (1, 1, 1, 1, 3)	1.44 1.16 (1, 1, 1, 2, 3)	1.58 1.30 (1, 1, 1, 2, 4)	1.75 1.47 (1, 1, 1, 2, 4)	1.94 1.65 (1, 1, 1, 2, 5)	2.16 1.85 (1, 1, 1, 3, 6)	2.40 2.07 (1, 1, 2, 3, 6)
7	1.44 1.16 (1, 1, 1, 2, 4)	1.60 1.32 (1, 1, 1, 2, 4)	1.78 1.50 (1, 1, 1, 2, 5)	1.98 1.70 (1, 1, 1, 2, 5)	2.21 1.92 (1, 1, 1, 3, 6)	2.48 2.17 (1, 1, 2, 3, 7)	2.78 2.46 (1, 1, 2, 3, 7)
9	1.60 1.33 (1, 1, 1, 2, 4)	1.79 1.52 (1, 1, 1, 2, 5)	2.01 1.73 (1, 1, 1, 2, 5)	2.26 1.97 (1, 1, 1, 3, 6)	2.54 2.25 (1, 1, 2, 3, 7)	2.88 2.57 (1, 1, 2, 4, 8)	3.25 2.93 (1, 1, 2, 4, 9)
11	1.80 1.53 (1, 1, 1, 2, 5)	2.03 1.76 (1, 1, 1, 2, 5)	2.29 2.02 (1, 1, 1, 3, 6)	2.60 2.32 (1, 1, 2, 3, 7)	2.96 2.67 (1, 1, 2, 4, 8)	3.37 3.06 (1, 1, 2, 4, 9)	3.85 3.53 (1, 1, 3, 5, 11)
13	2.03 1.77 (1, 1, 1, 2, 5)	2.31 2.04 (1, 1, 1, 3, 6)	2.63 2.36 (1, 1, 2, 3, 7)	3.02 2.74 (1, 1, 2, 4, 8)	3.46 3.18 (1, 1, 2, 4, 10)	3.99 3.69 (1, 1, 3, 5, 11)	4.62 4.30 (1, 1, 3, 6, 13)
15	2.32 2.05 (1, 1, 1, 3, 6)	2.66 2.39 (1, 1, 2, 3, 7)	3.06 2.80 (1, 1, 2, 4, 8)	3.53 3.26 (1, 1, 2, 5, 10)	4.10 3.82 (1, 1, 3, 5, 11)	4.79 4.50 (1, 1, 3, 6, 14)	5.60 5.28 (1, 2, 4, 7, 16)
17	2.67 2.41 (1, 1, 2, 3, 7)	3.09 2.83 (1, 1, 2, 4, 8)	3.59 3.32 (1, 1, 2, 5, 10)	4.19 3.92 (1, 1, 3, 5, 12)	4.93 4.65 (1, 1, 3, 6, 14)	5.80 5.51 (1, 2, 4, 8, 17)	6.87 6.56 (1, 2, 5, 9, 20)
19	3.11 2.84 (1, 1, 2, 4, 9)	3.62 3.36 (1, 1, 2, 5, 10)	4.25 3.99 (1, 1, 3, 6, 12)	5.04 4.77 (1, 1, 3, 7, 14)	5.97 5.69 (1, 2, 4, 8, 17)	7.12 6.83 (1, 2, 5, 10, 21)	8.56 8.25 (1, 2, 6, 11, 25)
21	3.63 3.37 (1, 1, 2, 5, 10)	4.29 4.03 (1, 1, 3, 6, 12)	5.11 4.85 (1, 1, 3, 7, 15)	6.09 5.82 (1, 2, 4, 8, 17)	7.32 7.04 (1, 2, 50, 10, 21)	8.87 8.58 (1, 3, 6, 12, 26)	10.80 10.50 (1, 3, 7, 15, 32)
23	4.31 4.05 (1, 1, 3, 6, 12)	5.15 4.90 (1, 1, 3, 7, 15)	6.17 5.91 (1, 2, 4, 8, 18)	7.47 7.20 (1, 2, 5, 10, 22)	9.10 8.83 (1, 3, 6, 12, 26)	11.18 10.90 (1, 3, 8, 15, 33)	14.08 13.78 (1, 4, 10, 19, 41)

* The three rows of each cell shows the ARL_0^+, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.

† See SAS Program 7 in Appendix B for the calculation of the values in Table 3.29.

Table 3.29 continued for $h = 2, 4, ..., 14$ and $k = 25, 27, ..., 53$.

k	h						
	2	**4**	**6**	**8**	**10**	**12**	**14**
25	5.17 4.92 (1, 1, 3, 7, 15)	6.22 5.97 (1, 2, 4, 8, 18)	7.56 7.30 (1, 2, 5, 10, 22)	9.28 9.01 (1, 3, 6, 13, 27)	11.47 11.19 (1, 3, 8, 16, 34)	14.57 14.29 (1, 4, 10, 20, 43)	18.65 18.35 (1, 5, 13, 25, 55)
27	6.25 5.99 (1, 2, 4, 8, 18)	7.62 7.36 (1, 2, 5, 10, 22)	9.39 9.13 (1, 3, 6, 13, 27)	11.67 11.40 (1, 3, 8, 16, 34)	14.94 14.67 (1, 4, 10, 20, 44)	19.28 19.00 (1, 5, 13, 26, 57)	24.76 24.47 (1, 7, 17, 34, 73)
29	7.64 7.39 (1, 2, 5, 10, 22)	9.46 9.20 (1, 3, 6, 13, 28)	11.80 11.54 (1, 3, 8, 16, 35)	15.20 14.94 (1, 4, 10, 21, 45)	19.75 19.48 (1, 6, 14, 27, 58)	25.55 25.27 (1, 7, 18, 35, 76)	33.84 33.55 (2, 10, 23, 47, 101)
31	9.48 8.97 (1, 3, 7, 13, 27)	11.87 11.36 (1, 4, 8, 16, 35)	15.37 14.85 (1, 5, 11, 21, 45)	20.08 19.56 (2, 6, 14, 28, 59)	26.12 25.58 (2, 8, 18, 36, 77)	34.85 34.29 (2, 10, 24, 48, 103)	47.67 47.38 (2, 14, 33, 66, 142)
33	11.91 11.40 (1, 4, 8, 16, 35)	15.47 14.96 (1, 5, 11, 21, 45)	20.30 19.79 (2, 6, 14, 28, 60)	26.51 25.99 (2, 8, 19, 37, 78)	35.57 35.04 (2, 11, 25, 49, 105)	49.00 48.45 (3, 14, 34, 68, 146)	68.60 68.32 (4, 20, 47, 95, 205)
35	15.52 15.01 (1, 5, 11, 21, 45)	20.43 19.92 (2, 6, 14, 28, 60)	26.76 26.25 (2, 8, 19, 37, 79)	36.06 35.54 (2, 11, 25, 50, 107)	49.95 49.42 (3, 15, 35, 69, 149)	70.38 69.83 (4, 21, 49, 97, 210)	97.26 96.98 (5, 28, 67, 134, 291)
37	20.48 19.97 (2, 6, 14, 28, 60)	26.90 26.39 (2, 8, 19, 37, 80)	36.35 35.85 (2, 11, 25, 50, 108)	50.56 50.04 (3, 15, 35, 70, 150)	71.58 71.05 (4, 21, 50, 99, 213)	99.34 98.80 (6, 29, 69, 138, 297)	161.57 161.29 (8, 46, 112, 224, 483)
39	26.95 26.44 (2, 8, 19, 37, 80)	36.51 36.00 (2, 11, 25, 50, 108)	50.92 50.41 (3, 15, 35, 70, 152)	72.34 71.83 (4, 21, 50, 100, 216)	100.71 100.19 (6, 29, 70, 139, 301)	165.16 164.62 (9, 48, 115, 229, 494)	245.13 244.86 (13, 70, 170, 339, 734)
41	36.57 36.07 (2, 11, 26, 51, 109)	51.12 50.62 (3, 15, 36, 71, 152)	72.81 72.30 (4, 21, 51, 101, 217)	101.58 101.07 (6, 30, 71, 141, 303)	167.59 167.07 (9, 49, 116, 232, 501)	249.70 249.18 (13, 72, 173, 346, 747)	486.87 486.60 (25, 140, 337, 675, 1458)
43	51.20 50.70 (3, 15, 36, 71, 152)	73.05 72.55 (4, 21, 51, 101, 218)	102.08 101.58 (6, 30, 71, 141, 305)	169.12 168.61 (9, 49, 117, 234, 506)	252.64 252.13 (13, 73, 175, 350, 756)	497.29 496.76 (26, 143, 345, 689, 1489)	
45	73.14 72.64 (4, 21, 51, 101, 218)	102.32 101.82 (6, 30, 71, 142, 306)	170.00 169.50 (9, 49, 118, 235, 508)	254.38 253.87 (14, 74, 176, 352, 761)	504.08 503.56 (26, 145, 350, 699, 1509)		
47	102.40 101.90 (6, 30, 71, 142, 306)	170.44 169.94 (9, 49, 118, 236, 510)	255.38 254.87 (14, 74, 177, 354, 764)	508.02 507.51 (27, 147, 352, 704, 1521)			
49	170.67 170.17 (9, 49, 118, 236, 510)	255.87 255.37 (14, 74, 178, 355, 766)	510.50 510.00 (27, 147, 354, 708, 1528)				
51	256.00 255.50 (14, 74, 178, 355, 766)	511.50 511.00 (27, 148, 355, 709, 1531)					
53	512.00 511.50 (27, 148, 355, 710, 1533)						

Table 3.29 continued for h = 16, 18, ..., 28 and k = 1, 3, ..., 25.

k	***h***						
	16	**18**	**20**	**22**	**24**	**26**	**28**
1	2.02 1.67 (1, 1, 1, 2, 5)	2.21 1.84 (1, 1, 1, 3, 6)	2.42 2.02 (1, 1, 2, 3, 6)	2.64 2.21 (1, 1, 2, 3, 7)	2.88 2.41 (1, 1, 2, 4, 7)	3.14 2.63 (1, 1, 2, 4, 8)	3.42 2.86 (1, 1, 2, 4, 9)
3	2.31 1.96 (1, 1, 1, 3, 6)	2.55 2.17 (1, 1, 2, 3, 7)	2.80 2.40 (1, 1, 2, 3, 7)	3.09 2.66 (1, 1, 2, 4, 8)	3.40 2.93 (1, 1, 2, 4, 9)	3.74 3.22 (1, 1, 3, 5, 10)	4.10 3.54 (1, 1, 3, 5, 11)
5	2.67 2.32 (1, 1, 2, 3, 7)	2.97 2.59 (1, 1, 2, 4, 8)	3.30 2.90 (1, 1, 2, 4, 9)	3.67 3.24 (1, 1, 2, 5, 10)	4.07 3.60 (1, 1, 3, 5, 11)	4.52 4.00 (1, 1, 3, 6, 12)	5.00 4.44 (1, 2, 3, 6, 14)
7	3.12 2.77 (1, 1, 2, 4, 8)	3.50 3.13 (1, 1, 2, 4, 9)	3.93 3.53 (1, 1, 3, 5, 11)	4.41 3.98 (1, 1, 3, 6, 12)	4.95 4.48 (1, 2, 3, 6, 14)	5.55 5.03 (1, 2, 4, 7, 15)	6.22 5.65 (1, 2, 4, 8, 17)
9	3.68 3.34 (1, 1, 2, 5, 10)	4.19 3.82 (1, 1, 3, 5, 12)	4.75 4.35 (1, 1, 3, 6, 13)	5.39 4.95 (1, 2, 4, 7, 15)	6.12 5.64 (1, 2, 4, 8, 17)	6.94 6.43 (1, 2, 5, 9, 20)	7.92 7.35 (1, 2, 5, 11, 22)
11	4.42 4.08 (1, 1, 3, 6, 12)	5.06 4.69 (1, 1, 3, 7, 14)	5.81 5.42 (1, 2, 4, 8, 16)	6.68 6.25 (1, 2, 5, 9, 19)	7.69 7.22 (1, 2, 5, 10, 22)	8.91 8.39 (1, 3, 6, 12, 25)	10.31 9.74 (1, 3, 7, 14, 29)
13	5.35 5.01 (1, 2, 4, 7, 15)	6.21 5.85 (1, 2, 4, 8, 18)	7.23 6.84 (1, 2, 5, 10, 21)	8.43 8.00 (1, 2, 6, 11, 24)	9.93 9.46 (1, 3, 7, 13, 29)	11.68 11.16 (1, 3, 8, 16, 34)	13.66 13.10 (1, 4, 9, 18, 40)
15	6.57 6.23 (1, 2, 4, 9, 19)	7.73 7.37 (1, 2, 5, 10, 22)	9.13 8.74 (1, 3, 6, 12, 26)	10.93 10.50 (1, 3, 7, 15, 32)	13.08 12.61 (1, 4, 9, 18, 38)	15.56 15.05 (1, 5, 11, 21, 45)	18.62 18.05 (1, 6, 13, 25, 54)
17	8.18 7.85 (1, 2, 6, 11, 24)	9.77 9.42 (1, 3, 7, 13, 28)	11.87 11.49 (1, 3, 8, 16, 35)	14.44 14.02 (1, 4, 10, 20, 42)	17.46 17.00 (1, 5, 12, 24, 51)	21.30 20.79 (1, 6, 15, 29, 63)	26.12 25.56 (2, 8, 18, 36, 77)
19	10.34 10.01 (1, 3, 7, 14, 30)	12.73 12.37 (1, 4, 9, 17, 37)	15.72 15.34 (1, 5, 11, 21, 46)	19.30 18.89 (1, 6, 13, 26, 57)	23.97 23.51 (1, 7, 17, 33, 71)	30.02 29.52 (2, 9, 21, 41, 89)	37.67 37.12 (2, 11, 26, 52, 111)
21	13.47 13.14 (1, 4, 9, 18, 39)	16.86 16.51 (1, 5, 12, 23, 50)	21.00 20.62 (1, 6, 14, 29, 62)	26.51 26.10 (1, 8, 18, 36, 78)	33.87 33.42 (2, 10, 23, 47, 100)	43.45 42.95 (2, 13, 30, 60, 129)	55.02 54.49 (3, 16, 38, 76, 163)
23	17.85 17.52 (1, 5, 12, 24, 53)	22.49 22.15 (1, 6, 15, 31, 66)	28.83 28.50 (2, 8, 20, 40, 85)	37.51 37.11 (2, 11, 26, 52, 111)	49.11 48.67 (3, 14, 34, 68, 146)	63.40 62.93 (3, 18, 44, 87, 189)	87.75 87.21 (5, 25, 61, 121, 262)
25	23.75 23.44 (1, 7, 16, 33, 70)	30.85 30.51 (2, 9, 21, 42, 91)	40.78 40.42 (2, 12, 28, 56, 121)	54.39 54.00 (3, 16, 38, 75, 162)	71.44 71.02 (4, 21, 49, 99, 213)	102.25 101.77 (5, 30, 71, 141, 305)	137.75 137.23 (7, 40, 95, 191, 411)

Table 3.29 continued for h = 16, 18, ..., 28 and k = 27, 29, ..., 53.

k	***h***						
	16	**18**	**20**	**22**	**24**	**26**	**28**
27	32.52 32.21 (2, 9, 22, 45, 97)	43.58 43.25 (2, 13, 30, 60, 130)	59.06 58.70 (3, 17, 41, 81, 176)	78.74 78.36 (4, 23, 54, 109, 235)	116.24 115.81 (6, 33, 80, 161, 347)	160.22 159.76 (8, 46, 111, 222, 479)	250.03 249.48 (13, 72, 173, 346, 748)
29	45.88 45.57 (2, 13, 32, 63, 137)	63.00 62.68 (3, 18, 44, 87, 188)	85.05 84.70 (4, 24, 59, 118, 254)	129.08 128.69 (7, 37, 79, 179, 386)	181.52 181.11 (9, 52, 126, 251, 543)	297.24 296.75 (15, 86, 206, 412, 889)	
31	66.17 65.87 (3, 19, 46, 91, 197)	90.22 89.91 (5, 26, 62, 125, 269)	140.19 139.84 (7, 40, 97, 194, 419)	200.46 200.10 (10, 58, 139, 278, 600)	343.38 342.95 (18, 99, 238, 476, 1028)		
33	94.26 93.97 (5, 27, 65, 130, 282)	149.28 148.96 (8, 43, 103, 207, 446)	216.42 216.09 (11, 62, 150, 300, 647)	385.45 385.06 (20, 111, 267, 534, 1154)			
35	156.39 156.09 (8, 45, 108, 216, 468)	229.10 228.79 (12, 66, 159, 317, 685)	412.40 421.06 (22, 121, 292, 584, 1261)				
37	238.47 238.18 (12, 69, 165, 330, 714)	450.00 449.68 (23, 129, 312, 623, 1347)					
39	471.21 471.12 (24, 136, 327, 653, 1411)						
41							
43							
45							
47							
49							
51							
53							

Table 3.29 continued for $h = 30, 32, ..., 42$ and $k = 1, 3, ..., 53$.

k	h						
	30	**32**	**34**	**36**	**38**	**40**	**42**
1	3.71 3.11 (1, 1, 3, 5, 10)	4.02 3.36 (1, 1, 3, 5, 10)	4.34 3.63 (1, 2, 3, 6, 11)	4.69 3.92 (1, 2, 3, 6, 12)	5.05 4.22 (1, 2, 4, 6, 13)	5.41 4.53 (1, 2, 4, 7, 14)	5.81 4.85 (1, 2, 4, 7, 15)
3	4.40 3.88 (1, 1, 3, 6, 12)	4.90 4.24 (1, 2, 3, 6, 13)	5.36 4.63 (1, 2, 4, 7, 14)	5.83 5.05 (1, 2, 4, 8, 16)	6.33 5.48 (1, 2, 4, 8, 17)	6.86 5.95 (1, 2, 5, 9, 18)	7.41 6.44 (1, 3, 5, 10, 20)
5	5.53 4.91 (1, 2, 4, 7, 15)	6.12 5.45 (1, 2, 4, 8, 17)	6.75 6.02 (1, 2, 5, 9, 18)	7.42 6.63 (1, 2, 5, 10, 20)	8.14 7.29 (1, 3, 6, 11, 22)	8.92 8.00 (1, 3, 6, 12, 25)	9.75 8.75 (1, 3, 7, 13, 27)
7	6.98 6.36 (1, 2, 5, 9, 19)	7.82 7.13 (1, 3, 5, 10, 22)	8.71 7.97 (1, 3, 6, 12, 24)	9.70 8.90 (1, 3, 7, 13, 27)	10.79 9.92 (1, 3, 8, 14, 30)	11.96 11.02 (1, 4, 8, 16, 34)	13.22 12.21 (1, 4, 9, 18, 37)
9	9.01 8.38 (1, 3, 6, 12, 25)	10.19 9.51 (1, 3, 7, 14, 29)	11.54 10.79 (1, 4, 8, 15, 33)	13.05 12.24 (1, 4, 9, 18, 37)	14.72 13.85 (1, 5, 10, 20, 42)	16.53 15.59 (1, 5, 12, 22, 47)	18.65 17.63 (2, 6, 13, 25, 54)
11	11.86 11.23 (1, 4, 8, 16, 34)	13.66 12.98 (1, 4, 9, 18, 39)	15.74 15.00 (1, 5, 11, 21, 45)	18.10 17.29 (1, 6, 13, 25, 52)	20.69 19.81 (2, 6, 14, 28, 60)	23.85 22.90 (2, 7, 17, 32, 69)	27.20 26.17 (2, 8, 19, 37, 79)
13	16.04 15.42 (1, 5, 11, 22, 47)	18.86 18.18 (1, 6, 13, 26, 55)	22.14 21.29 (2, 7, 15, 30, 65)	25.79 24.99 (2, 8, 18, 35, 75)	30.47 29.58 (2, 9, 21, 42, 89)	35.45 34.50 (2, 11, 25, 49, 104)	41.67 40.64 (3, 12, 29, 57, 123)
15	22.35 21.73 (1, 7, 15, 30, 65)	26.81 26.13 (2, 8, 19, 37, 79)	31.89 31.15 (2, 9, 22, 44, 94)	38.71 37.89 (3, 11, 27, 53, 114)	46.05 45.17 (3, 14, 32, 63, 136)	55.68 54.71 (4, 16, 39, 77, 165)	
17	32.04 31.43 (2, 9, 22, 44, 95)	38.92 38.26 (2, 11, 27, 53, 115)	48.70 47.95 (3, 14, 34, 67, 144)	59.32 58.51 (4, 17, 41, 82, 176)	74.09 73.19 (4, 22, 51, 102, 220)		
19	46.74 46.14 (3, 14, 32, 64, 139)	60.40 59.72 (3, 18, 42, 83, 179)	75.44 74.70 (4, 22, 52, 104, 224)	97.76 96.94 (6, 28, 68, 135, 291)			
21	73.57 72.96 (4, 21, 51, 102, 219)	94.27 93.61 (5, 27, 65, 130, 281)	127.36 126.60 (7, 37 , 88, 176, 380)				
23	115.36 114.77 (6, 33, 80, 159, 344)	163.12 162.43 (9, 47, 113, 226, 487)					
25	204.54 203.92 (11, 59, 142, 283, 611)						
27							
29							
⋮							
53							

Table 3.29 continued for $h = 44, 46, ..., 54$ and $k = 1, 3, ..., 53$.

k	***h***					
	44	**46**	**48**	**50**	**52**	**54**
1	6.21 5.19 (1, 2, 4, 8, 16)	6.63 5.53 (1, 2, 5, 9, 17)	7.05 5.89 (1, 3, 5, 9, 19)	7.50 6.27 (1, 3, 5, 10, 20)	7.95 6.65 (1, 3, 6, 10, 21)	8.43 7.05 (1, 3, 6, 11, 22)
3	8.00 6.95 (1, 3, 6, 10, 22)	8.60 7.48 (1, 3, 6, 11, 23)	9.25 8.06 (1, 3, 7, 12, 25)	9.91 8.65 (1, 4, 7, 13, 27)	10.62 9.28 (1, 4, 8, 14, 29)	
5	10.61 9.55 (1, 4, 7, 14, 29)	11.57 10.43 (1, 4, 8, 15, 32)	12.55 11.35 (1, 4, 9, 17, 35)	13.63 12.34 (2, 5, 10, 18, 38)		
7	14.64 13.56 (1, 5, 10, 20, 41)	16.11 14.96 (2, 5, 11, 22, 46)	17.77 16.54 (2, 6, 12, 24, 50)			
9	20.88 19.79 (2, 7, 15, 28, 60)	33.44 22.28 (2, 7, 16, 32, 68)				
11	31.19 30.09 (2, 10, 22, 43, 91)					
13						
15						
⋮						
53						

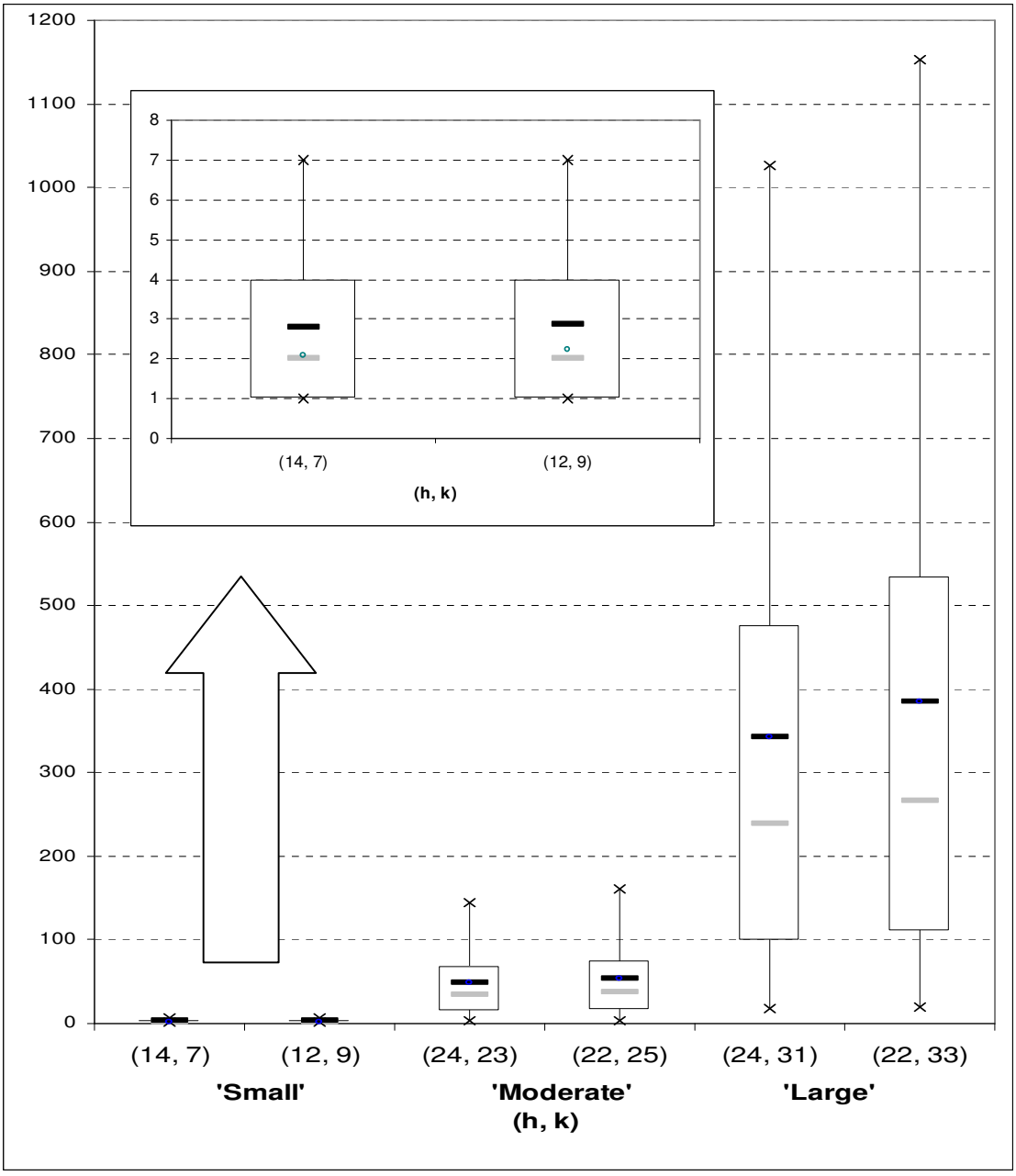

Figure 3.11. Boxplot-like graphs for the in-control run length distribution of various two-sided CUSUM signed-rank charts when $n = 10$. The whiskers extend to the 5th and the 95th percentiles. The symbols "▬", "◇" and "▬" denote the *ARL*, *SDRL*[*] and *MRL*, respectively[†].

[*] For ease of interpretation, the standard deviation (as measure of spread) is included in the (location) measures of percentiles.

[†] The "boxplots" are classified into 3 categories, namely small ($h+k \leq 25$), moderate ($25 < h+k \leq 50$) and large ($h+k > 50$).

Example 3.9

A two-sided CUSUM signed-rank chart for the Montgomery (2001) piston ring data

We conclude this sub-section by illustrating the two-sided CUSUM signed-rank chart using the piston ring data set from Montgomery (2001). We assume that the underlying distribution is symmetric with a known target value of $\theta_0 = 74\,mm$. For illustration take $k = 3$ and $h = 8$. From Table 3.27 it can be seen that the in-control average run length equals 4.07 when $(h,k) = (8,3)$. Generally, one chooses the chart constants so that a specified in-control average run length, such as 500, or 370, is obtained. Taking this into consideration, an in-control average run length of 4.07 is considered small. Table 3.30 shows the upper and lower signed-rank CUSUM statistics, respectively.

Table 3.30. One-sided signed-rank (S_i^+ and S_i^-) statistics*.

Sample number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
S_i^+	5	6	0	4	0	6	13	4	13	24	25	37	49	61	72
S_i^-	0	0	-11	-1	-1	0	0	-3	0	0	0	0	0	0	0

* The values in Table 3.30 we generated using Microsoft Excel.

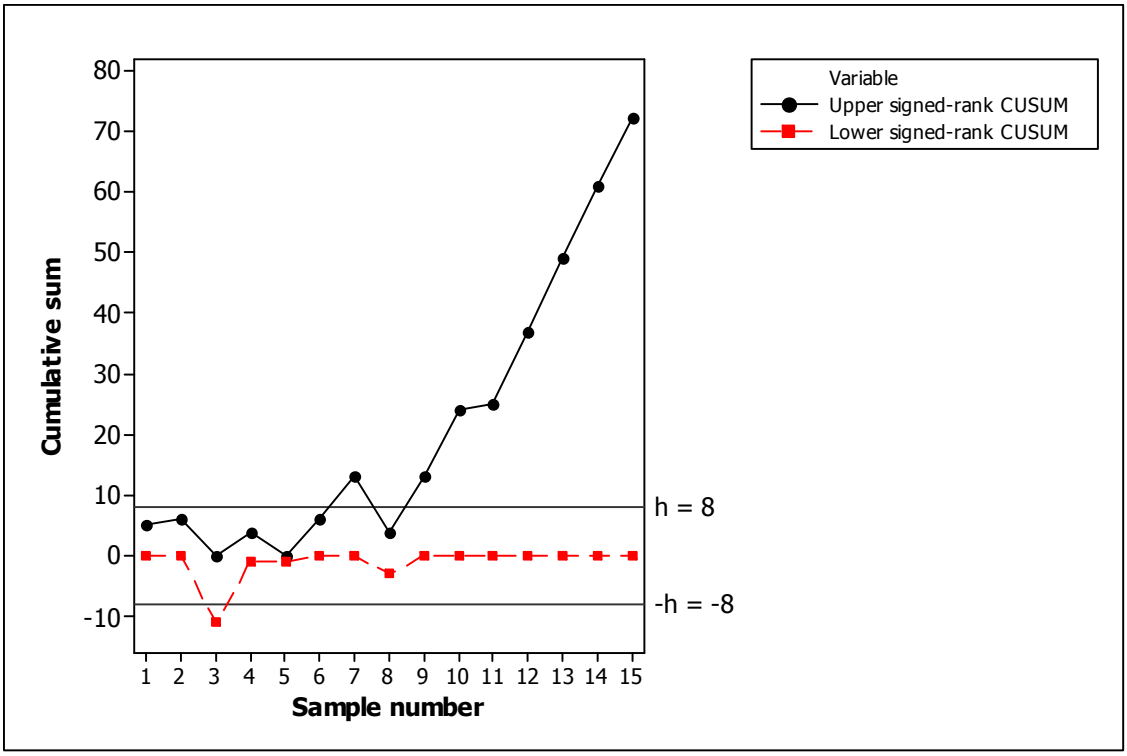

Figure 3.12. The two-sided CUSUM signed-rank chart for the Montgomery (2001) piston ring data.

The two-sided CUSUM signed-rank chart signals at sample number 3, indicating a most likely upward shift in the process median. The action taken following an out-of-control signal on a CUSUM chart is identical to that with any control chart. A search for assignable causes should be done, corrective action should be taken (if required) and, following this, the CUSUM is reset to zero.

3.3.4. Summary

While the Shewhart-type charts are the most widely known and used control charts 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. In this chapter we have described the properties of the CUSUM signed-rank chart and given tables for its implementation. Detailed calculations have been given to help the reader to understand the subject more thoroughly.

3.4. The EWMA control chart

3.4.1. Introduction

In this section, the approach taken by Lucas and Saccucci (1990) is extended to the use of the signed-rank statistic resulting in an EWMA signed-rank chart that accumulates the statistics $SR_1, SR_2, SR_3, \ldots$. Section 3.4 is analogous to Section 2.4 where the approach taken by Lucas and Saccucci (1990) was extended to the use of the sign statistic resulting in an EWMA sign chart. Therefore, the reader is frequently referred back to Section 2.4 throughout this section.

3.4.2. The proposed EWMA signed-rank chart

A nonparametric EWMA-type of control chart based on the signed-rank statistic (recall that $SR_i = \sum_{j=1}^{n} sign(x_{ij} - \theta_0) R_{ij}^{+}$) can be obtained by replacing X_i in expression (2.53) of Section 2.4 with SR_i. The EWMA signed-rank chart accumulates the statistics $SR_1, SR_2, SR_3, \ldots$ with the plotting statistics defined as

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

where $0 < \lambda \leq 1$ is a constant called the weighting constant. The starting value Z_0 could be taken to equal zero, i.e. $Z_0 = 0$.

The EWMA signed-rank chart is constructed by plotting Z_i against the sample number i (or time). If the plotting statistic Z_i falls between the two control limits, that is, $LCL < Z_i < UCL$, the process is considered to be in-control. If the plotting 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 out-of-control.

The exact control limits and the center line of the EWMA signed-rank control chart can be obtained by replacing σ and θ_0 by σ_{SR_i} and 0, respectively, in expression (2.55) of Section 2.4 to obtain

$$\begin{aligned} UCL &= L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \\ CL &= 0 \\ LCL &= -L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)} \end{aligned} \quad . \tag{3.11}$$

Similarly, the *steady-state* control limits can be obtained by replacing σ and θ_0 by σ_{SR_i} and 0, respectively, in expression (2.56) to obtain

$$\begin{aligned} UCL &= L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \\ LCL &= -L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)} \end{aligned} \tag{3.12}$$

where σ_{SR_i} denotes the in-control standard deviation of the signed-rank statistic SR_i if there are no ties within a subgroup.

The in-control standard deviation of SR_i is given by $\sigma_{SR_i} = \sqrt{\text{var}(SR_i)} = \sqrt{\text{var}\left(2T^+ - \frac{n(n+1)}{2}\right)} = \sqrt{\frac{n(n+1)(2n+1)}{6}}$. This is obtained by using the relationship between SR_i and T^+ (recall that $SR_i = 2T^+ - \frac{n(n+1)}{2}$ if there are no ties within a subgroup) and the fact that $\text{var}(T^+) = \frac{n(n+1)(2n+1)}{24}$ (see Gibbons and Chakraborti (2003) page 198).

3.4.3. Markov-chain approach

Lucas and Saccucci (1990) evaluated the properties of the *continuous state* Markov chain by *discretizing* the infinite state TPM. This procedure entails dividing the interval between the UCL and the LCL into N subintervals of width 2δ. Then the plotting statistic, Z_i, is said to be in the non-absorbing state j at time i if $S_j - \delta < Z_i \leq S_j + \delta$ where S_j denotes the midpoint of the j^{th} interval. Z_i is said to be in the absorbing state if Z_i falls on or outside one of the control

limits, that is, $Z_i \le LCL$ or $Z_i \ge UCL$. Let p_{ij} denote the probability of moving from state i to state j in one step, i.e. $p_{ij} = P(Moving\ to\ state\ j \mid in\ state\ i)$. To approximate this probability we assume that the plotting statistic is equal to S_i whenever it is in state i. For all j non-absorbing we obtain $p_{ij} = P(S_j - \delta < Z_k \le S_j + \delta \mid Z_{k-1} = S_i)$. By using the definition of the plotting statistic given in expression (3.10) we obtain

$$\begin{aligned} p_{ij} &= P(S_j - \delta < \lambda SR_k + (1-\lambda)S_i \le S_j + \delta) \\ &= P\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} < SR_k \le \frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda}\right) \end{aligned}$$

recall that $SR_k = 2T_k^+ - \frac{n(n+1)}{2}$

$$\begin{aligned} p_{ij} &= P\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} < 2T_k^+ - \frac{n(n+1)}{2} \le \frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(\frac{\left(\frac{(S_j - \delta) - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}\right)}{2} < T_k^+ \le \frac{\left(\frac{(S_j + \delta) - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}\right)}{2}\right). \end{aligned} \tag{3.13}$$

For all j absorbing we obtain

$$\begin{aligned} p_{ij} &= P(Z_k \le LCL \mid Z_{k-1} = S_i) + P(Z_k \ge UCL \mid Z_{k-1} = S_i) \\ &= P(\lambda SR_k + (1-\lambda)S_i \le LCL) + P(\lambda SR_k + (1-\lambda)S_i \ge UCL) \\ &= P\left(SR_k \le \frac{LCL - (1-\lambda)S_i}{\lambda}\right) + P\left(SR_k \ge \frac{UCL - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(2T_k^+ - \frac{n(n+1)}{2} \le \frac{LCL - (1-\lambda)S_i}{\lambda}\right) + P\left(2T_k^+ - \frac{n(n+1)}{2} \ge \frac{UCL - (1-\lambda)S_i}{\lambda}\right) \\ &= P\left(T_k^+ \le \frac{\frac{LCL - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}}{2}\right) + P\left(T_k^+ \ge \frac{\frac{UCL - (1-\lambda)S_i}{\lambda} + \frac{n(n+1)}{2}}{2}\right). \end{aligned} \tag{3.14}$$

Since the values LCL, UCL, δ, λ, n, S_i and S_j are known constants the Wilcoxon signed-rank probabilities in expressions (3.13) and (3.14) can easily be calculated. The probabilities for the Wilcoxon signed-rank statistics are given in Table H of Lehmann (1975) for

samples sizes up to 20 and they are tabulated (more recently) in Table H of Gibbons and Chakraborti (2003) for sample sizes up to 15.

Once the one-step transition probabilities are calculated, the TPM can be constructed and is given by $TPM = [p_{ij}] = \begin{pmatrix} Q & | & \underline{p} \\ - & - & - \\ \underline{0}' & | & 1 \end{pmatrix}$ (written in partitioned form) where the essential transition probability sub-matrix Q is the matrix that contains all the transition probabilities of going from a non-absorbing state to a non-absorbing state, $Q:(NA \rightarrow NA)$, $\underline{p}$ contains all the transition probabilities of going from each non-absorbing state to the absorbing states, $\underline{p}:(NA \rightarrow A)$, $\underline{0}' = (0 \quad 0 \quad 0 \quad \cdots \quad 0)$ contains all the transition probabilities of going from each absorbing state to the non-absorbing states. $\underline{0}'$ is a row vector with all its elements equal to zero, because it is impossible to go from an absorbing state to a non-absorbing state, because once an absorbing state is entered, it is never left, $\underline{0}' :(A \rightarrow NA)$, and 1 represents the scalar value one. The probability of going of going from an absorbing state to an absorbing state is equal to one, because once an absorbing state is entered, it is never left, $1:(A \rightarrow A)$. The one-step TPM is used to calculate the expected value (*ARL*), the second raw moment, the variance, the standard deviation and the probability mass function (pmf) of the run-length variable N which are given in equations (2.41) to (2.45).

Example 3.10

The EWMA signed-rank chart where the sample size is even ($n = 6$)

The EWMA signed-rank chart is investigated for a smoothing constant of 0.1 ($\lambda = 0.1$) and a multiplier of 3 ($L = 3$). The *steady-state* control limits are given by

$$UCL = L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

$$LCL = -L\sigma_{SR_i}\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

where $L = 3$, $\lambda = 0.1$, and $\sigma_{SR_i} = 9.539$, since $\sigma_{SR_i} = \sqrt{\frac{n(n+1)(2n+1)}{6}} = \sqrt{\frac{6(6+1)(12+1)}{6}} = 9.539$. Clearly, we only have to calculate the UCL since $LCL = -UCL$. We obtain $UCL = 3 \times 9.539\sqrt{\left(\frac{0.1}{2-0.1}\right)} = 6.565$. Therefore, $LCL = -6.565$.

This Markov-chain procedure entails dividing the interval between the UCL and the LCL into N subintervals of width 2δ. For this example N is taken to equal 4. Figure 3.13 illustrates the partitioning of the interval between the UCL and the LCL into subintervals.

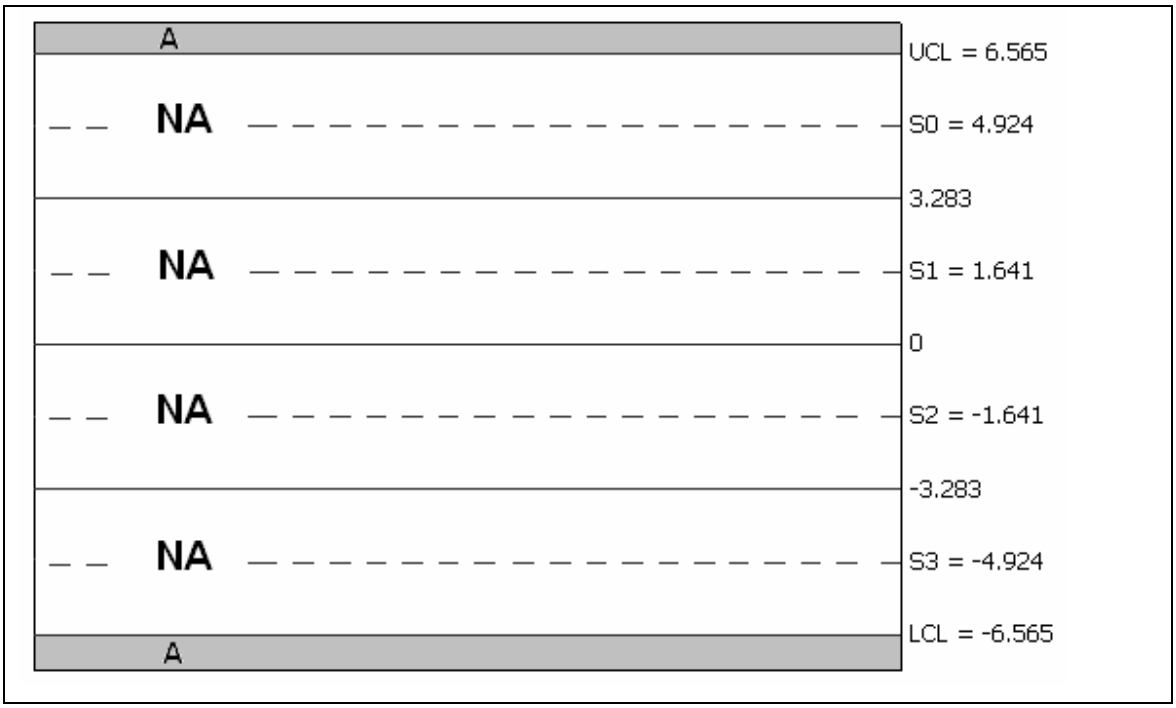

Figure 3.13. Partitioning of the interval between the UCL and the LCL into 4 subintervals.

From Figure 3.13 we see that there are 4 non-absorbing states, i.e. $r = 4$. The TPM is given by

$$TPM = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} \\ p_{10} & p_{11} & p_{12} & p_{13} & p_{14} \\ p_{20} & p_{21} & p_{22} & p_{23} & p_{24} \\ p_{30} & p_{31} & p_{32} & p_{33} & p_{34} \\ p_{40} & p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \begin{pmatrix} Q_{4\times4} & | & \underline{p}_{4\times1} \\ - & - & - \\ \underline{0}'_{1\times4} & | & 1_{1\times1} \end{pmatrix}.$$

Table 3.31. Calculation of the one-step probabilities in the first row of the TPM.

$p_{00} = P(\textit{Moving to state } 0 \mid \textit{in state } 0)$

$= P(S_0 - \delta < Z_k \le S_0 + \delta \mid Z_{k-1} = S_0)$ from (3.13)

$$= P\left(\frac{\left(\frac{(S_0 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_0 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$

with $\delta = 1.641$, $\lambda = 0.1$, $L = 3$ and $S_0 = 4.924$

$= P(4.755 < T_k^+ \le 21.169)$

$= P(T_k^+ \le 21) - P(T_k^+ \le 4)$

$= \frac{57}{64}$ from Gibbons and Chakraborti (2003)

$p_{01} = P(\textit{Moving to state } 1 \mid \textit{in state } 0)$

$= P(S_1 - \delta < Z_k \le S_1 + \delta \mid Z_{k-1} = S_0)$ from (3.13)

$$= P\left(\frac{\left(\frac{(S_1 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_1 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$

$= P(-11.658 < T_k^+ \le 4.755)$

$= P(T_k^+ \le 4)$

$= \frac{7}{64}$

$$p_{02} = P(\textit{Moving to state } 2 \mid \textit{in state } 0)$$
$$= P(S_2 - \delta < Z_k \le S_2 + \delta \mid Z_{k-1} = S_0) \text{ from (3.13)}$$
$$= P\left(\frac{\left(\frac{(S_2 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_2 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$
$$= P(-28.072 < T_k^+ \le -11.658)$$
$$= 0$$

$$p_{03} = P(\textit{Moving to state } 3 \mid \textit{in state } 0)$$
$$= P(S_3 - \delta < Z_k \le S_3 + \delta \mid Z_{k-1} = S_0) \text{ from (3.13)}$$
$$= P\left(\frac{\left(\frac{(S_3 - \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} < T_k^+ \le \frac{\left(\frac{(S_3 + \delta) - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2} \right)}{2} \right)$$
$$= P(-44.486 < T_k^+ \le -28.072)$$
$$= 0$$

$$p_{04} = P(\textit{Moving to state } 4 \mid \textit{in state } 0)$$
$$= P(Z_k \le LCL \mid Z_{k-1} = S_0) + P(Z_k \ge UCL \mid Z_{k-1} = S_0) \text{ from (3.14)}$$
$$= P\left(T_k^+ \le \frac{\frac{LCL - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2}}{2} \right) + P\left(T_k^+ \ge \frac{\frac{UCL - (1-\lambda)S_0}{\lambda} + \frac{n(n+1)}{2}}{2} \right)$$
$$= P(T_k^+ \le -44.486) + P(T_k^+ \ge 21.169)$$
$$= 0$$

The one-step probabilities in the remaining rows can be calculated similarly. Therefore, the TPM is given by $TPM = \begin{pmatrix} 57/64 & 7/64 & 0 & 0 & 0 \\ 2/64 & 57/64 & 5/64 & 0 & 0 \\ 0 & 5/64 & 57/64 & 2/64 & 0 \\ 0 & 0 & 7/64 & 57/64 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_{4\times4} & | & \underline{p}_{4\times1} \\ - & - & - \\ \underline{0}'_{1\times4} & | & 1_{1\times1} \end{pmatrix}$.

Other values of the multiplier (L) and the smoothing constant (λ) were also considered and the results are given in Tables 3.32 and 3.33.

Table 3.32. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values* for the EWMA signed-rank chart when $n = 6$ and $N = 5$, i.e. there are 5 subintervals between the lower and upper control limit†.

	$L = 1$	$L = 2$	$L = 3$
$\lambda = 0.05$	10.45 12.32 (1, 2, 6, 15, 35)	56.69 72.45 (1, 5, 29, 82, 204)	**
$\lambda = 0.1$	7.32 8.38 (1, 1, 4, 10, 24)	33.83 40.28 (1, 4, 20, 48, 115)	330.67 369.33 (2, 63, 213, 471, 1070)
$\lambda = 0.2$	4.95 4.90 (1, 1, 3, 7, 15)	35.21 39.63 (1, 6, 22, 50, 115)	361.92 384.29 (3, 87, 243, 510, 1130)

** The inverse of the matrix $(I - Q)$ does not exist and as a result the ARL (given by $E(N) = \underline{\xi}(I - Q)^{-1}\underline{1}$) can not be calculated for this combination of (λ, L).

In example 3.10 we considered a sample size that may be considered "small". The results are given for a larger sample size ($n = 10$) for various values of λ and L in Table 3.33.

Table 3.33. The in-control average run length (ARL_0), standard deviation of the run length ($SDRL$), 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentile values‡ for the EWMA signed-rank chart when $n = 10$ and $N = 5$, i.e. there are 5 subintervals between the lower and upper control limit§.

	$L = 1$	$L = 2$	$L = 3$
$\lambda = 0.05$	11.17 13.49 (1, 2, 6, 16, 39)	67.94 83.82 (1, 7, 38, 98, 238)	1448.44 1573.37 (10, 316, 956, 2052, 4595)
$\lambda = 0.1$	6.85 7.74 (1, 1, 4, 9, 23)	48.87 57.73 (1, 6, 29, 70, 165)	352.72 384.51 (3, 76, 232, 500, 1122)
$\lambda = 0.2$	5.05 5.07 (1, 1, 3, 7, 15)	33.96 38.48 (1, 6, 21, 48, 111)	336.34 357.54 (3, 80, 226, 474, 1051)

* The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.
† See SAS Program 8 in Appendix B for the calculation of the values in Table 3.32.
‡ The three rows of each cell shows the ARL_0, the $SDRL$, and the percentiles (ρ_5, ρ_{25}, ρ_{50}, ρ_{75}, ρ_{95}), respectively.
§ See SAS Program 8 in Appendix B for the calculation of the values in Table 3.33.

These tables can be extended by changing the sample size (n), the number of subintervals between the lower and upper control limit (N), the multiplier (L) and the smoothing constant (λ) in SAS Program 8 for the EWMA signed-rank chart given in Appendix B.

From Tables 3.32 and 3.33 we see that the ARL_0, $SDRL$ and percentiles increase as the value of the multiplier (L) increases. From Table 3.33 we find an in-control average run length of 336.34 for $n = 10$ when the multiplier is taken to equal 3 ($L = 3$) and the smoothing constant 0.2 ($\lambda = 0.2$). The chart performance is good, since the attained in-control average run length of 336.34 is in the region of the desired in-control average run length which is generally taken to be 370 or 500.

3.4.4. Summary

The EWMA control chart is one of several charting methods aimed at correcting a deficiency of the Shewhart chart - insensitivity to small shifts. Lucas and Saccucci (1990) have investigated some properties of the EWMA chart under the assumption of independent normally distributed observations, whereas in this section we have described and evaluated the nonparametric EWMA signed-rank chart. The main advantage of the nonparametric EWMA chart is that there is no need to assume a particular parametric distribution for the underlying process (see Section 1.4 for other advantages of the nonparametric EWMA chart).

Section B: Monitoring the location of a process when the target location is unspecified or unknown (Case U)

Introduction

In Section A we focussed on monitoring the location of a chart when the location is specified (case K). This 'standard(s) known' case is when the underlying parameters of the process distribution are known or specified. In Section B we focus on monitoring the location of a chart when the location is unspecified or unknown (case U). This 'standard(s) unknown' case is when the parameters are unknown and need to be estimated.

Chapter 4: Sign-like control charts

4.1. The Shewhart-type control chart

4.1.1. Introduction

Janacek and Meikle (1997) proposed a Phase II nonparametric control chart useful in case U. The control limits of this chart are given by two selected order statistics of a Phase I reference sample. The charting statistic is the median M_i of the Phase II samples taken sequentially.

Chakraborti, Van der Laan and Van de Wiel (2004; hereafter CVV) generalized the work of Janacek and Meikle (1997). They considered using some order statistic of a Phase II sample as the charting statistic and control limits constructed from a Phase I reference sample. Their work involves a class of two-sample nonparametric statistics, called precedence statistics and their Shewhart-type charts are called precedence charts. The terms *precedence charts* and *sign-like charts* will be used interchangeably throughout this text.

Assume that a reference sample of size m, $X_1, X_2, ..., X_m$, is available from an in-control process with an unknown continuous cdf $F(x)$. The estimated control limits of the precedence chart are given by two reference sample order statistics, say, $L\hat{C}L = X_{(a:m)}$ and

$\hat{UCL} = X_{(b:m)}$, where $1 \le a < b \le m$. Let $Y_1^h, Y_2^h, ..., Y_{n_h}^h$, $h = 1,2,...,$ denote the h^{th} test sample of size n_h. The plotting statistic $Y_{(j:n_h)}^h$ is the j^{th} order statistic from the h^{th} Phase II sample of size n_h. Let $G^h(y)$ denote the cdf of the distribution of the h^{th} Phase II sample. $G^h(y) = G(y)$ $\forall h$, since the Phase II samples are all assumed to be identically distributed. Assume that the Phase II samples are all of the same size, n, so that the subscript h can be suppressed. Under this assumption the plotting statistic is denoted by $Y_{(j:n)}$. For illustration purposes the plotting statistic is taken to be the median, but it can be any percentile of the Phase II sample. CVV provided recommendations and tables for the implementation of precedence charts and examined the chart performance in terms of the average run length. The overall conclusion is that the Shewhart-type precedence charts are more robust than their parametric counterparts, such as the Shewhart $\overline{X}$ chart. The precedence chart, being nonparametric, has the in-control robustness property (such as the same ARL_0 or the *FAR* for all continuous distributions), whereas as we noted earlier, the performance of the Shewhart $\overline{X}$ (and other parametric charts) is significantly (highly) degraded if the distributional form of the observations differs from normality.

4.1.2. Preliminary

Let W_j denote the number of X-observations that precede $Y_{(j:n)}$. The statistic W_j is called a precedence statistic and subsequently a test based on a precedence statistic is called a precedence test. Chakraborti and Van der Laan (1996, 1997; hereafter CV) gave an overview of some nonparametric procedures based on precedence statistics. CV's procedures included both hypothesis testing and confidence intervals. CV also highlighted the fact that precedence tests are simple and robust nonparametric procedures that are useful for comparing two or more distributions.

Let $P_C(W_j = w)$ denote the in-control probability distribution of W_j, where the subscript C refers to the in-control case. If $W_j = w$ it means that w X-observations precede $Y_{(j:n)}$. If w X-observations are less than or equal to $Y_{(j:n)}$, then $(m - w)$ X-observations are greater than $Y_{(j:n)}$. If we combine the reference sample (containing m X-observations) with

the test sample (containing n Y-observations) we obtain a single sample consisting of $N = m + n$ observations. From this combined sample, w X-observations and j Y-observations are less than or equal to $Y_{(j:n)}$. On the other hand, $(m - w)$ X-observations and $(n - j)$ Y-observations are greater than $Y_{(j:n)}$. There are a total of $w + j - 1$ observations that are less than $Y_{(j:n)}$ and a total of $(m - w) + (n - j) = m + n - j - w$ observations that are greater than $Y_{(j:n)}$. The in-control distribution of W_j can be obtained by using combinatorics which allows one to count the number of experimental outcomes when the experiment involves selecting a number of objects, say r, from a larger set of objects, say R. The rule then states that the number of combinations of R objects taken r at a time is given by $\binom{R}{r}$.

By using such combinatorial arguments the in-control distribution of W_j can be obtained and is given by

$$P_C(W_j = w) = \frac{\binom{w + j - 1}{w}\binom{m + n - j - w}{m - w}}{\binom{m + n}{m}} \quad \text{with} \quad w = 0,1,2,...,m. \qquad (4.1)$$

Note that the in-control probability distribution of W_j, i.e. when $F = G$, only depends on the number of observations in the reference sample m, the number of observations in each test sample n and the chosen percentile of the Phase II sample j. Thus, the in-control run length distribution of these precedence charts are distribution-free. The only condition is that the distribution of the reference sample and the distribution of the test sample be continuous and identical which is the case when the process is under control. It should be noted that this result is also given by Randles and Wolfe (1979), Theorem 11.4.4.

As illustration, let the number of observations in the Phase I reference sample be 25 ($m = 25$), the number of observations in each Phase II test sample be 15 ($n = 15$) and the chosen percentile of the Phase II sample be the median $\left(j = \frac{n+1}{2} = \frac{15+1}{2} = 8\right)$. The in-control distribution, i.e. when $F = G$, of W_j is then given by

$$P_C(W_j = w) = \frac{\binom{w+7}{w}\binom{32-w}{25-w}}{\binom{40}{25}} \quad \text{with} \quad w = 0,1,...,25. \tag{4.2}$$

Figure 4.1 represents the in-control distribution of W_j when $m = 25$, $n = 15$ and $j = 8$. Note that the in-control probability distribution of W_j is symmetric. In general, the in-control probability distribution of W_j is symmetric when n is odd and the chosen percentile of the Phase II sample is the median of an odd Phase II sample.

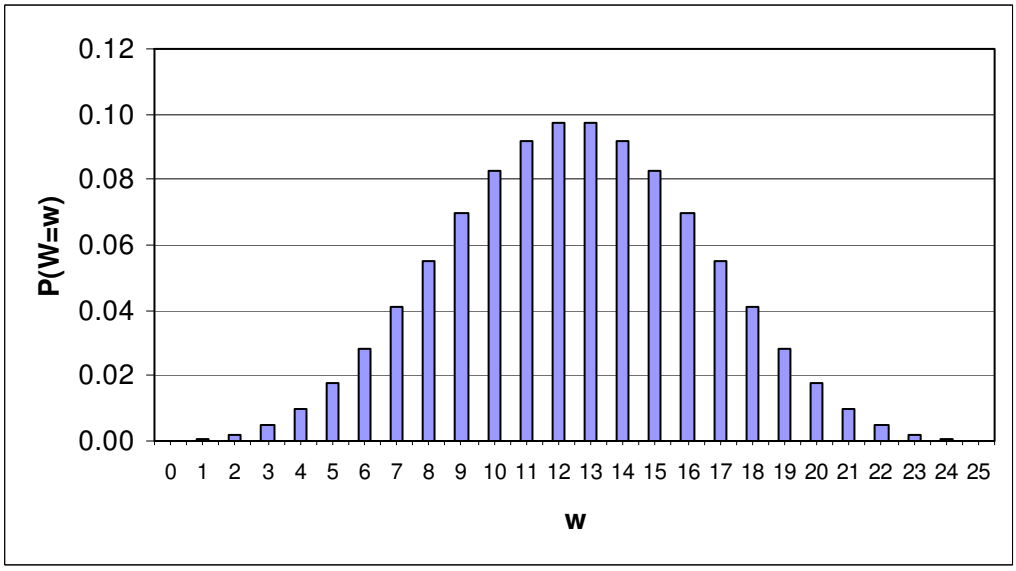

Figure 4.1. The in-control distribution of W_j when $m = 25$, $n = 15$ and $j = 8$.

Figure 4.2 represents the in-control probability distribution of W_j when the number of observations in the reference sample is kept at 25 ($m = 25$), the number of observations in each test sample is kept at 15 ($n = 15$), but the chosen percentile of the Phase II sample is not the median, i.e. $j \neq 8$. We take $j = 4$ for illustration purposes. Note that the in-control probability distribution of W_j is now asymmetric.

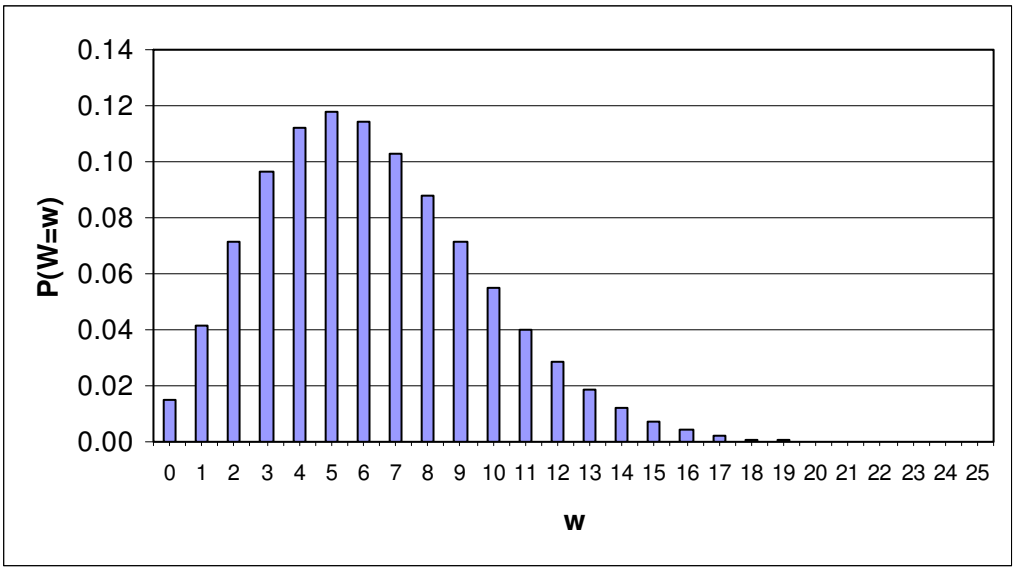

Figure 4.2. The in-control distribution of W_j when $m = 25$, $n = 15$ and $j = 4$.

4.1.3. Probability of no signal

Recall that $L\hat{C}L = X_{(a:m)}$ and $U\hat{C}L = X_{(b:m)}$. A non-signalling event in the case of the two-sided chart occurs when $X_{(a:m)} \le Y_{(j:n)} \le X_{(b:m)}$. Stated differently, a non-signalling event occurs when at least a X-observations precede $Y_{(j:n)}$ and at most $b-1$ X-observations precede $Y_{(j:n)}$, i.e. $a \le W_j \le b-1$. Let the probability of no signal be denoted by p. Then, the probability of no signal is given by

$$p = p(m,n,j;F,G) = P\left(X_{(a:m)} \le Y_{(j:n)} \le X_{(b:m)}\right) = P\left(a \le W_j \le b-1\right). \qquad (4.3)$$

From (4.3) it can be seen that the probability of no signal, p, can be expressed in terms of the precedence statistic W_j, thus simplifying the probability calculations (see Randles and Wolfe (1979), Example 11.4.19).

Let p_0 denote the in-control value of p. A process is said to be in-control when $G = F$. Therefore, the expression for p_0 can be obtained by simply substituting $G = F$ in expression (4.3). Thus,

$$p_0 = p(m,n,j;F,F) = P(\text{No Signal} \mid \text{In - control}) = P_C\left(a \le W_j \le b-1\right). \qquad (4.4)$$

Recall that a false alarm is given when a signaling event occurs, given that the process is actually in-control. Therefore, the probability of a false alarm (also referred to as the false alarm rate (*FAR*)) is given by

$$1-p_0 = 1-P(\text{No Signal} \mid \text{In-control}) = P(\text{Signal} \mid \text{In-control}) = FAR\,. \qquad (4.5)$$

4.1.4. Determination of chart constants

The charting constants a and b are typically selected so that a specified false alarm rate or a specified in-control average run-length is attained. The exact expression for the ARL_0 is derived later on in this chapter using a conditioning method. In this section we will focus on the *FAR*. Hence, the charting constants a and b are found by either setting the *FAR* (given by $1-p_0$) to a desirable small value, say $1-P_0$, or by setting p_0 to some desirable large value, say P_0. Take note that P_0 will usually be chosen to be a large value such as 0.95 or 0.99 and the desired or specified value of the *FAR*, given by $1-P_0$, will be a small value such as 0.05 or 0.01. The charting constants are found such that $p_0 = P(\text{No Signal} \mid \text{In-control})$ is not smaller than the desired or specified value P_0, that is, $p_0 = P(\text{No Signal} \mid \text{In-control}) \geq P_0$ (this is due to the discrete nature of the distribution of W_j). Stated differently, the charting constants are selected such that $1-p_0 = P(\text{Signal} \mid \text{In-control})$ is not larger than the desired or specified value $1-P_0$, that is, $1-p_0 = P(\text{Signal} \mid \text{In-control}) \leq 1-P_0$. Since the statistic W_j is discrete, not all desired or specified P_0 values are attainable for all combinations of m, n and j. The inequality sign in (4.6) ensures that we are conservative. The charting constants are found such that

$$P_C(a \leq W_j \leq b-1) = \sum_{w=a}^{b-1} \frac{\binom{w+j-1}{w}\binom{m+n-j-w}{m-w}}{\binom{m+n}{m}} \geq p_0\,. \qquad (4.6)$$

We can use any test sample order statistic (including the median) when implementing the two-sided precedence chart. If the plotting statistic is taken to be the median, the in-control probability distribution of W_j is symmetric (for odd sample sizes) and a reasonable choice for b is $m-a+1$. Once the charting constants a and b are found, the estimated

control limits $L\hat{C}L = X_{(a:m)}$ and $U\hat{C}L = X_{(b:m)}$ can be determined. Therefore, when the plotting statistic is taken to be the median, we replace b by $m-a+1$ in (4.6) to obtain

$$P_C(a \le W_j \le m-a) = \sum_{w=a}^{m-a} \frac{\binom{w+j-1}{w}\binom{m+n-j-w}{m-w}}{\binom{m+n}{m}} \ge p_0 . \qquad (4.7)$$

For example, let the number of observations in the reference sample be 125 ($m = 125$), the number of observations in each test sample be 5 ($n = 5$) and the chosen percentile of the Phase II sample be the median $\left(j = \frac{n+1}{2} = \frac{5+1}{2} = 3 \quad (\text{when } n \text{ is odd}) \right)$. By substituting $m = 125$, $n = 5$ and $j = 3$ into (4.7) we obtain

$$P_C(a \le W_3 \le 125-a) = \sum_{w=a}^{125-a} \frac{\binom{w+2}{w}\binom{127-w}{125-w}}{\binom{130}{125}} \ge p_0 . \qquad (4.8)$$

Possible control limits were calculated using (4.8) and are shown in Table 4.1.

Table 4.1[*]**.** False alarm rate (FAR) and chart constant (a) values for the Shewhart sign-like chart when $m = 125$, $n = 5$ and $j = 3$.

a	3	4	5	6	7	8	9	10
FAR	0.000546	0.001079	0.001865	0.002948	0.004368	0.006164	0.008372	0.011025

From Table 4.1 we see that for a false alarm rate of 0.004368 one can take $a = 7$ so $b = m - a + 1 = 125 - 7 + 1 = 119$ so that the control limits are the 7th and 119th ordered values of the reference sample. Thus, $L\hat{C}L = X_{(7:125)}$ and $U\hat{C}L = X_{(119:125)}$. For another example on exceedance statistics see Randles and Wolfe (1979), Example 11.4.19.

[*] The values in Table 4.1 were generated using Microsoft Excel. Table 4.1 is an extension of Table 3 given in Chakraborti, Eryilmaz and Human (2006).

4.1.5. The median chart

Let $n = 2s+1$, where $s = 0,1,2,...,$ (so that n is odd). Therefore, the median is uniquely given by $j = \frac{n+1}{2} = \frac{(2s+1)+1}{2} = s+1$. The statistic W_{s+1} is called the median statistic of Mathisen (1943). The in-control probability distribution of W_{s+1} is found by substituting $n = 2s+1$ and $j = s+1$ into (4.1) and is given by

$$P_C(W_{s+1} = w) = \frac{\binom{w+s}{w}\binom{m+s-w}{m-w}}{\binom{m+2s+1}{m}} \tag{4.9}$$

(see Randles and Wolfe (1979), Example 11.4.5). Recall that the in-control distribution of W_j (in this case, W_{s+1}) is symmetric when n is odd and the chosen percentile of the Phase II sample is the median. In this case a reasonable choice for b is $m-a+1$. The charting constant a is found by substituting $n = 2s+1$, $j = s+1$ and $b = m-a+1$ into equation (4.6) and then solving for a such that (4.10) is satisfied.

$$P_C(a \le W_{s+1} \le m-a) = \sum_{w=a}^{m-a} \frac{\binom{w+s}{w}\binom{m+s-w}{m-w}}{\binom{m+2s+1}{m}} \ge p_0. \tag{4.10}$$

Once the charting constant a is found using expression (4.10), the charting constant b is found from the relationship $b = m-a+1$. Thereafter, the control limits $L\hat{C}L = X_{(a:m)}$ and $U\hat{C}L = X_{(b:m)}$ can be determined. By using symmetry we have that

$$\begin{aligned} &P_C(a \le W_j \le m-a) \\ &= 1 - \left(P_C(0 \le W_j \le a-1) + P_C(m-a+1 \le W_j \le m)\right) \\ &= 1 - 2P_C(0 \le W_j \le a-1) \end{aligned}$$

and by setting $1 - 2P_C(0 \le W_j \le a-1) \ge P_0$ we obtain

$$P_C(0 \le W_j \le a-1) \le \frac{1-P_0}{2}. \tag{4.11}$$

Therefore, expression (4.10) can be re-written as expression (4.11) which is more convenient to work with.

For example, let the number of observations in the reference sample be 125 ($m = 125$) and $s = 2$ so that $n = 2s + 1 = 5$ and $j = s + 1 = 3$. By substituting $m = 125$ and $s = 2$ into (4.10) we obtain

$$P_C(a \le W_3 \le 125 - a) = \sum_{w=a}^{125-a} \frac{\binom{w+2}{w}\binom{127-w}{125-w}}{\binom{130}{125}} \ge p_0$$

which is equal to expression (4.8). Therefore, the FAR values given in Table 4.1 can be used in this example, meaning that one can take $a = 7$ so $b = m - a + 1 = 125 - 7 + 1 = 119$ so that the control limits are the 7^{th} and 119^{th} ordered values of the reference sample. Thus, $\hat{LCL} = X_{(7:125)}$ and $\hat{UCL} = X_{(119:125)}$, which yield a FAR of 0.004368.

4.1.6. Control charts for other percentiles

Since we could be interested in other percentiles than the median (see Radson and Boyd (2005) and Shmueli and Cohen (2003)), the distribution of W_j is not symmetric (in such cases) and finding the charting constants a and b is much more difficult.

Chakraborti, Van der Laan and Van de Wiel (2004) proposed the equal-tailed* procedure when the $100\rho^{th}$ percentile is of interest where $0 < \rho < 1$. The equal-tailed procedure is as follows:

Find the *largest* integer a $(1 \le a \le [m\rho])$ such that

$$P_C(0 \le W_j \le a - 1) \le \frac{1 - P_0}{2},$$

and the *smallest* integer b $(a < b \le m)$ such that

$$P_C(b + 1 \le W_j \le m) \le \frac{1 - P_0}{2}.$$

These a and b values are then substituted in the control limits $\hat{LCL} = X_{(a:m)}$ and $\hat{UCL} = X_{(b:m)}$.

* Note that in general $b \ne m - a + 1$ in this case so that the "equal-tailed" means equality in tail probabilities.

4.1.7. Properties of order statistics

The ordered values of a sample are known as the order statistics. Various authors have studied order statistics (see for example Randles and Wolfe (1979)). Our goal is to study the distribution of order statistics. In addition, we give some well-known properties and results of order statistics that will be used later on.

Suppose that $X_1, X_2, ..., X_n$ denotes a random sample of size n from a continuous pdf, $f(x)$. The pdf of the k^{th} order statistic $X_{(k:n)}$ is given by

$$g_k(x_{(k:n)}) = \frac{n!}{(k-1)!(n-k)!}\left(F(x_{(k:n)})\right)^{k-1}\left(1-F(x_{(k:n)})\right)^{n-k} f(x_{(k:n)}). \qquad (4.12)$$

The joint pdf for $X_{(k:n)}$ and $X_{(l:n)}$ is given by

$$g_{kl}(x_{(k:n)}, x_{(l:n)}) = \frac{n!}{(k-1)!(l-k-1)!(n-l)!}\times$$
$$\left(F(x_{(k:n)})\right)^{k-1} f(x_{(k:n)})\left(F(x_{(l:n)} - x_{(k:n)})\right)^{l-k-1}\left(1-F(x_{(l:n)})\right)^{n-l} f(x_{(l:n)}). \qquad (4.13)$$

Let $U_{(k:n)}$ denote the k^{th} order statistic of a sample of size n from the Uniform(0,1) distribution. The pdf of $U_{(k:n)}$ is given by

$$f_{U_{(k:n)}}(u) = \frac{1}{\beta(k, n-k+1)} u^{k-1}(1-u)^{n-k} \qquad (4.14)$$

where $\beta(k, n-k+1) = \dfrac{\Gamma(k)\Gamma(n-k+1)}{\Gamma(n+1)} = \dfrac{(k-1)!(n-k)!}{n!}$.

The binomial series arises in connection with distributions of order statistics. The binomial theorem gives the expansion of $(a+b)^k$. Using the binomial theorem we obtain

$$(a-b)^k = \sum_{n=0}^{k} (-1)^n \binom{k}{n} a^{k-n} b^n \qquad (4.15)$$

where a and b are any real numbers and k is a positive integer.

4.1.8. One-sided control charts

In this section the lower- and upper one-sided precedence control charts are considered. The lower one-sided chart will have a $L\hat{C}L$ equal to some constant value and an $U\hat{C}L = \infty$. In contrast, the upper one-sided chart will have an $U\hat{C}L$ equal to some constant and a $L\hat{C}L = -\infty$.

4.1.8.1. Lower one-sided control charts

For the lower one-sided chart we have the $L\hat{C}L = X_{(a:m)}$. Therefore, a non-signalling event occurs when $Y_{(j:n)} \geq X_{(a:m)}$.

Result 4.1: Probability of no signal - conditional

$$P\left(\text{No Signal} \mid X_{(a:m)} = x\right) = \int_{G(x)}^{1} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du$$

Using the probability integral transformation (PIT) (see, for example, Gibbons and Chakraborti (2003)), we know that $Y_{(j:n)} = G^{-1}(U_{(j:n)})$ and $X_{(a:m)} = F^{-1}(U_{(a:m)})$ where F and G are both continuous cdf's.

$$\begin{aligned}
&P\left(\text{No Signal} \mid X_{(a:m)} = x\right) = p_L(x), \text{ say,} \\
&= P\left(Y_{(j:n)} \geq x \mid X_{(a:m)} = x\right) \\
&= P\left(G^{-1}(U_{(j:n)}) \geq x \mid X_{(a:m)} = x\right) \\
&= P\left(U_{(j:n)} \geq G(x) \mid X_{(a:m)} = x\right) \\
&= \int_{G(x)}^{1} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du
\end{aligned}$$

since $\frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}$ is the pdf of $U_{(j:n)}$ (see equation (4.14)).

Result 4.2: Probability of no signal – unconditional

Let p_L denote the unconditional probability of no signal, then:

$$P(\text{No Signal}) = P\left(Y_{(j:n)} \geq X_{(a:m)}\right)$$

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(1 - GF^{-1}(v)\right)^{j+h} \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

$$p_L$$
$$= P(\text{No Signal})$$
$$= P\left(Y_{(j:n)} \geq X_{(a:m)}\right)$$
$$= E_{X_{(a:m)}}\left(P(Y_{(j:n)} \geq X_{(a:m)} \mid X_{(a:m)})\right)$$
$$= E_{X_{(a:m)}}\left(P(G(Y_{(j:n)}) \geq G(X_{(a:m)}) \mid X_{(a:m)})\right)$$

By the PIT we have that $U_{(j:n)} = G(Y_{(j:n)})$ where G is the continuous cdf of the Phase II sample $Y_1, Y_2, ..., Y_n$. Using this we obtain

$$= E_{X_{(a:m)}}\left(P(U_{(j:n)} \geq G(X_{(a:m)}) \mid X_{(a:m)})\right)$$

By the PIT we have that $U_{(a:m)} = F(X_{(a:m)})$ so that $X_{(a:m)} = F^{-1}(U_{(a:m)})$ where F is the continuous cdf of the reference sample $X_1, X_2, ..., X_m$. Using this we obtain

$$= E_{U_{(a:m)}}\left(P(U_{(j:n)} \geq GF^{-1}(U_{(a:m)}) \mid U_{(a:m)})\right)$$

$$= \int_0^1 P(U_{(j:n)} \geq GF^{-1}(v) \mid U_{(a:m)} = v) f(v) dv$$

where $f(v)$ is the pdf of $U_{(a:m)}$ which is given by $f(v) = \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}$

$$= \int_0^1 \left(\int_{GF^{-1}(v)}^1 \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du \right) f(v) dv$$

$$= \int_0^1 \left(\int_{GF^{-1}(v)}^1 \frac{1}{\beta(j, n-j+1)} u^{j-1} (1-u)^{n-j} \, du \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} \, dv$$

The term $(1-u)^{n-j}$ can be expanded to $(1-u)^{n-j} = \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} 1^{n-j-h} u^h =$ $\sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} u^h$ by using a binomial expansion (see equation (4.15)) and we obtain

$$= \int_0^1 \left(\int_{GF^{-1}(v)}^1 \frac{1}{\beta(j, n-j+1)} u^{j-1} \left(\sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} u^h \right) du \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} \, dv$$

By taking all the constants out of the integral sign and simplifying by setting $u^{j-1} u^h = u^{j-1+h}$ we obtain

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} \int_{GF^{-1}(v)}^1 u^{j-1+h} \, du \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} \, dv$$

Integrating $\int_{GF^{-1}(v)}^1 u^{j-1+h} du = \frac{u^{j+h}}{j+h}\Big|_{u=GF^{-1}(v)}^{u=1} = \frac{1^{j+h}}{j+h} - \frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h} = \frac{1-\left(GF^{-1}(v)\right)^{j+h}}{j+h}$ we have

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} \left(\frac{1-\left(GF^{-1}(v)\right)^{j+h}}{j+h} \right) \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} \, dv$$

which simplifies to

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(1 - GF^{-1}(v)\right)^{j+h} \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} \, dv \, .$$

Result 4.3: Probability of a signal - conditional

A signalling event occurs when $Y_{(j:n)} < X_{(a:m)}$.

$$P\left(\text{Signal} \mid X_{(a:m)} = x\right) = 1 - \int_{G(x)}^{1} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du$$

Result 4.4: Probability of a signal - unconditional

$$1 - p_L = P(\text{Signal}) =$$

$$1 - \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(1 - GF^{-1}(v)\right)^{j+h} \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

Result 4.5: Probability of a false alarm - conditional

$$CFAR = 1 - \int_{F(x)}^{1} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du$$

Follows immediately from Result 4.3, since $G = F$.

Result 4.6: Probability of a false alarm – unconditional

$$FAR = 1 - \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (1-v)^{j+h} \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

Follows immediately from Result 4.4, since $G = F$ and therefore $GF^{-1}(v) = FF^{-1}(v) = v$.

Result 4.7: Run-length distribution - conditional

$$P\left(N = k \mid X_{(a:m)} = x\right) = (p_L(x))^{k-1}(1 - p_L(x)) \quad \text{for} \quad k = 1,2,3,...$$

The conditional run length, denoted by $N \mid X_{(a:m)} = x$, will have a geometric distribution with parameter $1 - p_L(x)$, because all the signalling events are independent. Therefore we have that

$N \mid X_{(a:m)} = x \sim GEO(1 - p_L(x))$

$P\left(N = k \mid X_{(a:m)} = x\right) = (p_L(x))^{k-1}(1 - p_L(x)) \quad \text{for} \quad k = 1,2,3,...$

Consequently, the cumulative distribution function (cdf) is found from

$P\left(N \le k \mid X_{(a:m)} = x\right)$

$= \sum_{i=1}^{k} (p_L(x))^{i-1}(1 - p_L(x)) = 1 - (p_L(x))^k \quad \text{for} \quad k = 1,2,3,...$

We also have that

$P\left(N > k \mid X_{(a:m)} = x\right) = 1 - \left(1 - (p_L(x))^k\right) = (p_L(x))^k.$

Result 4.8: Average run-length - conditional

$$CARL = E\left(N \mid X_{(a:m)} = x\right) = \frac{1}{1 - p_L(x)}$$

or

$$CARL = E\left(N \mid X_{(a:m)} = x\right) = \sum_{k=0}^{\infty} (p_L(x))^k$$

Since the conditional run length, denoted by $N \mid X_{(a:m)} = x$ has a geometric distribution with parameter $1 - p_L(x)$, the conditional average run length is given by

$$CARL = E\left(N \mid X_{(a:m)} = x\right) = \frac{1}{1 - p_L(x)}.$$

The second expression follows immediately from the geometric expansion of $(1 - p_L(x))^{-1}$ for $p_L(x) < 1$.

Result 4.9: Run-length distribution - unconditional

$$P(N=k) = D_L^*(k-1) - D_L^*(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_L^*(0) = 1$$

where

$$D_L^*(k) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(1 - GF^{-1}(v)\right)^{j+h} \right)^k \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

$$\begin{aligned}
&P(N=k) \\
&= E_{X_{(a:m)}}\left(P\left(N=k \mid X_{(a:m)}=x\right)\right) \\
&= E_{X_{(a:m)}}\left((p_L(x))^{k-1}(1-p_L(x))\right) \\
&= E_{X_{(a:m)}}\left((p_L(x))^{k-1} - (p_L(x))^{k}\right) \\
&= E_{X_{(a:m)}}\left((p_L(x))^{k-1}\right) - E_{X_{(a:m)}}\left((p_L(x))^{k}\right)
\end{aligned}$$

By only focussing on $E_{X_{(a:m)}}\left((p_L(x))^k\right)$ we have

$$\begin{aligned}
&E_{X_{(a:m)}}\left((p_L(x))^k\right) \\
&= E_{U_{(a:m)}}\left(\left(\int_{GF^{-1}(v)}^1 \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du\right)^k\right) \\
&= \int_0^1 \left(\left(\int_{GF^{-1}(v)}^1 \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du\right)^k\right) f(v)\, dv
\end{aligned}$$

where $f(v)$ is the pdf of $U_{(a:m)}$ which is given by $f(v) = \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}$

$$= \int_0^1 \left(\left(\int_{GF^{-1}(v)}^1 \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du\right)^k\right) \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

The term $(1-u)^{n-j}$ can be expanded to $(1-u)^{n-j} = \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} 1^{n-j-h} u^h =$ $\sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} u^h$ by using a binomial expansion (see equation (4.15)) and we obtain

$$= \int_0^1 \left(\left(\int_{GF^{-1}(v)}^1 \frac{1}{\beta(j, n-j+1)} u^{j-1} \left(\sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} u^h \right) du \right)^k \right) \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} dv$$

By taking all the constants out of the integral sign and simplifying by setting $u^{j-1} u^h = u^{j-1+h}$ we obtain

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} \int_{GF^{-1}(v)}^1 u^{j-1+h} du \right)^k \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} dv$$

Integrating $\int_{GF^{-1}(v)}^1 u^{j-1+h} du = \frac{u^{j+h}}{j+h}\Big|_{u=GF^{-1}(v)}^{u=1} = \frac{1^{j+h}}{j+h} - \frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h} = \frac{1-\left(GF^{-1}(v)\right)^{j+h}}{j+h}$ we have

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} \left(\frac{1-\left(GF^{-1}(v)\right)^{j+h}}{j+h} \right) \right)^k \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} dv$$

which simplifies to

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(1 - GF^{-1}(v)\right)^{j+h} \right)^k \frac{m!}{(a-1)!(m-a)!} v^{a-1} (1-v)^{m-a} dv$$

$= D_L^*(k)$,
say.

Recall that $P(N=k) = E_{X_{(a:m)}}\left((p_L(x))^{k-1}\right) - E_{X_{(a:m)}}\left((p_L(x))^k\right)$.

Therefore, $P(N=k) = D_L^*(k-1) - D_L^*(k)$ for $k = 1,2,3,\ldots,$ and $D_L^*(0) = 1$.

$D_L^*(0) = 1$ since

$$D_L^*(0) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(1 - GF^{-1}(v)\right)^{j+h} \right)^0 \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

$$= \int_0^1 \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

$$= \int_0^1 f(v)dv$$

$$= 1.$$

This equals one, because in general, $\int_{-\infty}^{\infty} f(x)dx = 1$ for real x.

Result 4.10: In-control run-length distribution

$$P_C(N = k) = D_L(k-1) - D_L(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_L(0) = 1$$

and

$$D_L(k) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (1-v)^{j+h} \right)^k \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a}\, dv$$

Recall that the reference sample of size m, $X_1, X_2, \ldots, X_m$, is available from an in-control process with a continuous cdf, $F(x)$. The plotting statistic $Y_{(j:n)}^h$ is the j^{th} order statistic from the h^{th} Phase II sample of size n_h. Let $G^h(y)$ denote the cdf of the distribution of the h^{th} Phase II sample. A process is said to be in-control at stage h when $G^h = F$. Assume that the Phase II samples are all of the same size, n, so that the subscript h can be suppressed. Therefore, a process is said to be in-control when $G = F$. Therefore, the in-control run length distribution is obtained by setting $G = F$ into the equation for the out-of-control run length distribution.

The out-of-control run length distribution for the lower one-sided chart is given by

$$P(N = k) = D_L^*(k-1) - D_L^*(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_L^*(0) = 1$$

and (4.16)

$$D_L^*(k)=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(1-GF^{-1}(v)\right)^{j+h}\right)^k\frac{m!}{(a-1)!(m-a)!}v^{a-1}(1-v)^{m-a}\,dv.$$

Therefore, the in-control run length distribution for the lower one-sided chart is obtained by setting $G=F$ into equation (4.16) and we obtain

$$P_C(N=k)=D_L(k-1)-D_L(k)\quad\text{for}\quad k=1,2,3,\ldots,\quad D_L(0)=1$$

and (4.17)

$$D_L(k)=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(1-v\right)^{j+h}\right)^k\frac{m!}{(a-1)!(m-a)!}v^{a-1}(1-v)^{m-a}\,dv\,.$$

Result 4.11: Out-of-control average run-length - unconditional

$$UARL_{L,\delta}=\int_0^1\frac{1}{1-S_L(v,j,n,F,G)}f(v)dv$$

with

$$S_L(v,j,n,F,G)=\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(1-GF^{-1}(v)\right)^{j+h}$$

Let $UARL_{L,\delta}$ denote the unconditional average run length, where δ refers to the out-of-control case. To derive an expression for the $UARL_{L,\delta}$, recall that

$$\begin{aligned}&E_{X_{(a:m)}}\left((p_L(x))^k\right)\\&=D_L^*(k)\\&=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(1-GF^{-1}(v)\right)^{j+h}\right)^k\frac{m!}{(a-1)!(m-a)!}v^{a-1}(1-v)^{m-a}\,dv\\&=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(1-GF^{-1}(v)\right)^{j+h}\right)^k f(v)dv\end{aligned}$$

For simplicity let $S_L(v,j,n,F,G)=\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(1-GF^{-1}(v)\right)^{j+h}$, therefore we obtain

$$= \int_0^1 \left(S_L(v, j, n, F, G)\right)^k f(v)dv$$

Finally, we have that (from the second expression in Result 4.8)

$$UARL_{L,\delta}$$
$$= \sum_{k=0}^{\infty} E_{X_{(a:m)}}\left((p_L(x))^k\right)$$
$$= \sum_{k=0}^{\infty} D_L^*(k)$$
$$= \sum_{k=0}^{\infty} \int_0^1 \left(S_L(v, j, n, F, G)\right)^k f(v)dv$$
$$= \int_0^1 \sum_{k=0}^{\infty} \left(S_L(v, j, n, F, G)\right)^k f(v)dv$$
$$= \int_0^1 \frac{1}{1 - S_L(v, j, n, F, G)} f(v)dv \text{ from the geometric expansion of } (1 - S_L(v, j, n, F, G))^{-1}.$$

Result 4.12: In-control average run-length - unconditional

$$UARL_{L,0} = \int_0^1 \left(\frac{1}{1 - S_L(v, j, n)}\right) \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a} dv$$

with

$$S_L(v, j, n) = \frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (1-v)^{j+h}$$

Let $UARL_{L,0}$ denote the unconditional average run length, where 0 refers to the in-control case. To derive an expression for the $UARL_{L,0}$, recall that the in-control run length distribution for the lower one-sided chart is given by

$$P_C(N = k) = D_L(k-1) - D_L(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_L(0) = 1$$

with

$$D_L(k) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (1-v)^{j+h}\right)^k \frac{m!}{(a-1)!(m-a)!} v^{a-1}(1-v)^{m-a} dv$$

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (1-v)^{j+h} \right)^k f(v) dv$$

For simplicity let $S_L(v, j, n) = \frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (1-v)^{j+h}$, therefore we obtain

$$D_L(k) = \int_0^1 \left(S_L(v, j, n)\right)^k f(v) dv$$

and, finally, we have that

$$\begin{aligned} & UARL_{L,0} \\ &= \sum_{k=0}^{\infty} D_L(k) \\ &= \sum_{k=0}^{\infty} \int_0^1 \left(S_L(v, j, n)\right)^k f(v) dv \\ &= \int_0^1 \sum_{k=0}^{\infty} \left(S_L(v, j, n)\right)^k f(v) dv \\ &= \int_0^1 \frac{1}{1 - S_L(v, j, n)} f(v) dv . \end{aligned}$$

4.1.8.2. Upper one-sided control charts

For the upper one-sided chart we have $\hat{UCL} = X_{(b:m)}$. Therefore, a non-signalling event occurs when $Y_{(j:n)} \leq X_{(b:m)}$.

Result 4.13: Probability of no signal - conditional

$$P\left(\text{No Signal} \mid X_{(b:m)} = z\right) = \int_0^{G(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1} (1-u)^{n-j} du$$

Using the PIT, we know that $Y_{(j:n)} = G^{-1}(U_{(j:n)})$ and $X_{(b:m)} = F^{-1}(U_{(b:m)})$ where F and G are both continuous cdf's.

$$P(\text{No Signal} \mid X_{(b:m)} = z) = p_U(z), \text{ say,}$$
$$= P(Y_{(j:n)} \le z \mid X_{(b:m)} = z)$$
$$= P(G^{-1}(U_{(j:n)}) \le z \mid X_{(b:m)} = z)$$
$$= P(U_{(j:n)} \le G(z) \mid X_{(b:m)} = z)$$
$$= \int_0^{G(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du$$

since $\frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}$ is the pdf of $U_{(j:n)}$ (see equation (4.14)).

Result 4.14: Probability of no signal – unconditional

Let p_U denote the unconditional probability of no signal, then:

$$P(\text{No Signal}) = P(Y_{(j:n)} \le X_{(b:m)})$$
$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(GF^{-1}(v)\right)^{j+h} \right) \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

$$p_U$$
$$= P(\text{No Signal})$$
$$= P(Y_{(j:n)} \le X_{(b:m)})$$
$$= E_{X_{(b:m)}}\left(P(Y_{(j:n)} \le X_{(b:m)} \mid X_{(b:m)})\right)$$
$$= E_{X_{(b:m)}}\left(P(G(Y_{(j:n)}) \le G(X_{(b:m)}) \mid X_{(b:m)})\right)$$

By the PIT we have that $U_{(j:n)} = G(Y_{(j:n)})$ where G is the continous cdf of the Phase II sample $Y_1, Y_2, ..., Y_n$. Using this we obtain

$$= E_{X_{(b:m)}}\left(P(U_{(j:n)} \le G(X_{(b:m)}) \mid X_{(b:m)})\right)$$

By the PIT we have that $U_{(b:m)} = F(X_{(b:m)})$ so that $X_{(b:m)} = F^{-1}(U_{(b:m)})$ where F is the continous cdf of the reference sample $X_1, X_2, ..., X_m$. Using this we obtain

$$= E_{U_{(b:m)}}\left(P(U_{(j:n)} \le GF^{-1}(U_{(b:m)}) \mid U_{(b:m)})\right)$$
$$= \int_0^1 P(U_{(j:n)} \le GF^{-1}(v) \mid U_{(b:m)} = v) f(v) dv$$

where $f(v)$ is the pdf of $U_{(b:m)}$ which is given by $f(v)=\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}$

$$=\int_0^1\left(\int_0^{GF^{-1}(v)}\frac{1}{\beta(j,n-j+1)}u^{j-1}(1-u)^{n-j}\,du\right)\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}\,dv$$

The term $(1-u)^{n-j}$ can be expanded to $(1-u)^{n-j}=\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}1^{n-j-h}u^h=\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}u^h$ by using a binomial expansion (see equation (4.15)) and we obtain

$$=\int_0^1\left(\int_0^{GF^{-1}(v)}\frac{1}{\beta(j,n-j+1)}u^{j-1}\left(\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}u^h\right)du\right)\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}\,dv$$

By taking all the constants out of the integral sign and simplifying by setting $u^{j-1}u^h=u^{j-1+h}$ we obtain

$$=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}\int_0^{GF^{-1}(v)}u^{j-1+h}\,du\right)\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}\,dv$$

By integrating $\int_0^{GF^{-1}(v)}u^{j-1+h}\,du=\frac{u^{j+h}}{j+h}\Big|_{u=0}^{u=GF^{-1}(v)}=\frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h}-0=\frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h}$ we obtain

$$=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}\frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h}\right)\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}\,dv$$

which simplifies to

$$=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(v)\right)^{j+h}\right)\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}\,dv\,.$$

Result 4.15: Probability of a signal - conditional

A signalling event occurs when $Y_{(j:n)} > X_{(b:m)}$.

$$P(\text{Signal} \mid X_{(b:m)} = z) = 1 - \int_0^{G(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du$$

Result 4.16: Probability of a signal – unconditional

$$1 - p_U = P(\text{Signal}) =$$

$$1 - \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(GF^{-1}(v)\right)^{j+h} \right) \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

Result 4.17: Probability of a false alarm - conditional

$$CFAR = 1 - \int_0^{F(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du$$

Follows immediately from Result 4.15, since $G = F$.

Result 4.18: Probability of a false alarm - unconditional

$$FAR = 1 - \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} v^{j+h} \right) \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

Follows immediately from Result 4.16, since $G = F$ and therefore $GF^{-1}(v) = FF^{-1}(v) = v$.

Result 4.19: Run-length distribution - conditional

$$P\left(N = k \mid X_{(b:m)} = z\right) = (p_U(z))^{k-1}(1 - p_U(z)) \quad \text{for} \quad k = 1,2,3,...$$

The conditional run length, denoted by $N \mid X_{(b:m)} = z$, will have a geometric distribution with parameter $1 - p_U(z)$, because all the signalling events are independent. Therefore we have that

$N \mid X_{(b:m)} = z \sim GEO(1 - p_U(z))$

$P\left(N = k \mid X_{(b:m)} = z\right) = (p_U(z))^{k-1}(1 - p_U(z)) \quad \text{for} \quad k = 1,2,3,...$

Consequently, the cumulative distribution function (cdf) is found from

$P\left(N \le k \mid X_{(b:m)} = z\right)$

$= \sum_{i=1}^{k} (p_U(z))^{i-1}(1 - p_U(z)) = 1 - (p_U(z))^k \quad \text{for} \quad k = 1,2,3,...$

We also have that

$P\left(N > k \mid X_{(b:m)} = z\right) = 1 - \left(1 - (p_U(z))^k\right) = (p_U(z))^k$.

Result 4.20: Average run-length - conditional

$$CARL = E\left(N \mid X_{(b:m)} = z\right) = \frac{1}{1 - p_U(z)}$$

or

$$CARL = E\left(N \mid X_{(b:m)} = z\right) = \sum_{k=0}^{\infty} (p_U(z))^k$$

Since the conditional run length, denoted by $N \mid X_{(b:m)} = z$ has a geometric distribution with parameter $1 - p_U(z)$, the conditional average run length is given by

$$CARL = E\left(N \mid X_{(b:m)} = z\right) = \frac{1}{1 - p_U(z)}.$$

The second expression follows immediately from the geometric expansion of $(1 - p_U(z))^{-1}$ for $p_U(z) < 1$.

Result 4.21: Run-length distribution - unconditional

$$P(N=k) = D_U^*(k-1) - D_U^*(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_U^*(0) = 1$$

and

$$D_U^*(k) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(GF^{-1}(v)\right)^{j+h} \right)^k \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

$$\begin{aligned}
&P(N=k) \\
&= E_{X_{(b:m)}}\left(P\left(N=k \mid X_{(b:m)}=z\right)\right) \\
&= E_{X_{(b:m)}}\left((p_U(z))^{k-1}(1-p_U(z))\right) \\
&= E_{X_{(b:m)}}\left((p_U(z))^{k-1} - (p_U(z))^{k}\right) \\
&= E_{X_{(b:m)}}\left((p_U(z))^{k-1}\right) - E_{X_{(b:m)}}\left((p_U(z))^{k}\right)
\end{aligned}$$

By only focussing on $E_{X_{(b:m)}}\left((p_U(z))^k\right)$ we have

$$\begin{aligned}
&E_{X_{(b:m)}}\left((p_U(z))^k\right) \\
&= E_{U_{(b:m)}}\left(\left(\int_0^{GF^{-1}(v)} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du\right)^k\right)
\end{aligned}$$

$$= \int_0^1 \left(\left(\int_0^{GF^{-1}(v)} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du\right)^k\right) f(v)\, dv$$

where $f(v)$ is the pdf of $U_{(b:m)}$ which is given by $f(v) = \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}$

$$= \int_0^1 \left(\left(\int_0^{GF^{-1}(v)} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du\right)^k\right) \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

The term $(1-u)^{n-j}$ can be expanded to $(1-u)^{n-j} = \sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}1^{n-j-h}u^h =$

$\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}u^h$ by using a binomial expansion (see equation 4.15)) and we obtain

$$= \int_0^1 \left(\left(\int_0^{GF^{-1}(v)} \frac{1}{\beta(j,n-j+1)} u^{j-1} \left(\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}u^h \right) du \right)^k \right) \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

By taking all the constants out of the integral sign and simplifying by setting $u^{j-1}u^h = u^{j-1+h}$ we obtain

$$= \int_0^1 \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h} \int_0^{GF^{-1}(v)} u^{j-1+h}\, du \right)^k \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

By integrating $\int_0^{GF^{-1}(v)} u^{j-1+h}\, du = \frac{u^{j+h}}{j+h}\Big|_{u=0}^{u=GF^{-1}(v)} = \frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h} - 0 = \frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h}$ we obtain

$$= \int_0^1 \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h} \frac{\left(GF^{-1}(v)\right)^{j+h}}{j+h} \right)^k \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

which simplifies to

$$= \int_0^1 \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(v)\right)^{j+h} \right)^k \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

$= D_U^*(k)$,

say.

Recall that $P(N=k) = E_{X_{(b:m)}}\left((p_U(z))^{k-1}\right) - E_{X_{(b:m)}}\left((p_U(z))^k\right)$.

Therefore, $P(N=k) = D_U^*(k-1) - D_U^*(k)$ for $k = 1,2,3,\ldots,$ and $D_U^*(0) = 1$.

$D_U^*(0) = 1$ since

$$D_U^*(0) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(GF^{-1}(v)\right)^{j+h} \right)^0 \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

$$= \int_0^1 \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

$$= \int_0^1 f(v)dv$$

$$= 1.$$

This equals one, because in general, $\int_{-\infty}^{\infty} f(x)dx = 1$ for real x.

Result 4.22: In-control run-length distribution

$$P_C(N = k) = D_U(k-1) - D_U(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_U(0) = 1$$

and

$$D_U(k) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (v)^{j+h} \right)^k \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b}\, dv$$

Recall that the reference sample of size m, $X_1, X_2, \ldots, X_m$, is available from an in-control process with a continuous cdf, $F(x)$. The plotting statistic $Y_{(j:n)}^h$ is the j^{th} order statistic from the h^{th} Phase II sample of size n_h. Let $G^h(y)$ denote the cdf of the distribution of the h^{th} Phase II sample. A process is said to be in-control at stage h when $G^h = F$. Assume that the Phase II samples are all of the same size, n, so that the subscript h can be suppressed. Therefore, a process is said to be in-control when $G = F$. Therefore, the in-control run length distribution is obtained by setting $G = F$ into the equation for the out-of-control run length distribution.

The out-of-control run length distribution for the upper one-sided chart is given by

$$P(N = k) = D_U^*(k-1) - D_U^*(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_U^*(0) = 1$$

and (4.18)

$$D_U^*(k)=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(v)\right)^{j+h}\right)^k\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}dv.$$

Therefore, the in-control run length distribution for the upper one-sided chart is obtained by setting $G=F$ into equation (4.18) and we obtain

$$P_C(N=k)=D_U(k-1)-D_U(k)\quad\text{for}\quad k=1,2,3,\ldots,\quad D_U(0)=1 \tag{4.19}$$

and

$$D_U(k)=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}(v)^{j+h}\right)^k\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}dv.$$

Result 4.23: Out-of-control average run-length - unconditional

$$UARL_{U,\delta}=\int_0^1\frac{1}{1-S_U(v,j,n,F,G)}f(v)dv$$

with

$$S_U(v,j,n,F,G)=\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(v)\right)^{j+h}$$

Let $UARL_{U,\delta}$ denote the unconditional average run length, where δ refers to the out-of-control case. To derive an expression for the $UARL_{U,\delta}$, recall that

$$\begin{aligned}
&E_{X_{(b:m)}}\left((p_U(z))^k\right)\\
&=D_U^*(k)\\
&=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(v)\right)^{j+h}\right)^k\frac{m!}{(b-1)!(m-b)!}v^{b-1}(1-v)^{m-b}dv\\
&=\int_0^1\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(v)\right)^{j+h}\right)^k f(v)dv
\end{aligned}$$

For simplicity let $S_U(v,j,n,F,G)=\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}\frac{(-1)^h}{j+h}\binom{n-j}{h}\left(GF^{-1}(v)\right)^{j+h}$, therefore we obtain

$$D_U^*(k) = \int_0^1 \left(S_U(v, j, n, F, G)\right)^k f(v)dv$$

Finally, we have that (from the second expression in Result 4.20)

$$\begin{aligned}
&UARL_{U,\delta} \\
&= \sum_{k=0}^{\infty} E_{X_{(b:m)}}\left((p_U(z))^k\right) \\
&= \sum_{k=0}^{\infty} D_U^*(k) \\
&= \sum_{k=0}^{\infty} \int_0^1 \left(S_U(v, j, n, F, G)\right)^k f(v)dv \\
&= \int_0^1 \sum_{k=0}^{\infty} \left(S_U(v, j, n, F, G)\right)^k f(v)dv \\
&= \int_0^1 \frac{1}{1 - S_U(v, j, n, F, G)} f(v)dv .
\end{aligned}$$

Result 4.24: In-control average run-length - unconditional

$$UARL_{U,0} = \int_0^1 \left(\frac{1}{1 - S_U(v, j, n)}\right) \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b} dv$$

with

$$S_U(v, j, n) = \frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (v)^{j+h}$$

Let $UARL_{U,0}$ denote the unconditional average run length, where 0 refers to the in-control case. To derive an expression for the $UARL_{U,0}$, recall that the in-control run length distribution for the upper one-sided chart is given by

$$P_C(N = k) = D_U(k-1) - D_U(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D_U(0) = 1$$

with

$$D_U(k) = \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (v)^{j+h}\right)^k \frac{m!}{(b-1)!(m-b)!} v^{b-1}(1-v)^{m-b} dv$$

$$= \int_0^1 \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (v)^{j+h} \right)^k f(v) dv .$$

For simplicity let $S_U(v, j, n) = \frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} (v)^{j+h}$, therefore we obtain

$$D_U(k) = \int_0^1 \left(S_U(v, j, n)\right)^k f(v) dv$$

Finally, we have that

$$\begin{aligned}
&UARL_{U,0} \\
&= \sum_{k=0}^{\infty} D_U(k) \\
&= \sum_{k=0}^{\infty} \int_0^1 \left(S_U(v, j, n)\right)^k f(v) dv \\
&= \int_0^1 \sum_{k=0}^{\infty} \left(S_U(v, j, n)\right)^k f(v) dv \\
&= \int_0^1 \frac{1}{1 - S_U(v, j, n)} f(v) dv .
\end{aligned}$$

4.1.9. Two-sided control charts

For the two-sided chart we have $L\hat{C}L = X_{(a:m)}$ and $U\hat{C}L = X_{(b:m)}$. Therefore, a non-signalling event occurs when $X_{(a:m)} \le Y_{(j:n)} \le X_{(b:m)}$.

Result 4.25: Probability of no signal - conditional

$$P\left(\text{No Signal} \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = \int_{G(x)}^{G(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1} (1-u)^{n-j} du$$

Using the PIT, we know that $Y_{(j:n)} = G^{-1}(U_{(j:n)})$, $X_{(a:m)} = F^{-1}(U_{(a:m)})$ and $X_{(b:m)} = F^{-1}(U_{(b:m)})$ where F and G are both continuous cdf's.

$$P(\text{No Signal} \mid X_{(a:m)} = x, X_{(b:m)} = z) = p(x,z), \text{ say,}$$
$$= P(x \le Y_{(j:n)} \le z \mid X_{(a:m)} = x, X_{(b:m)} = z)$$
$$= P(x \le G^{-1}(U_{(j:n)}) \le z \mid X_{(a:m)} = x, X_{(b:m)} = z)$$
$$= P(G(x) \le U_{(j:n)} \le G(z) \mid X_{(a:m)} = x, X_{(b:m)} = z)$$

$$= \int_{G(x)}^{G(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1}(1-u)^{n-j}\, du$$

since $\frac{1}{\beta(j,n-j+1)} u^{j-1}(1-u)^{n-j}$ is the pdf of $U_{(j:n)}$ (see equation 4.14).

Result 4.26: Probability of no signal – unconditional

Let p denote the unconditional probability of no signal, then:

$$P(\text{No Signal}) = P(X_{(a:m)} \le Y_{(j:n)} \le X_{(b:m)})$$
$$\int_0^1 \int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right) \right) \times$$
$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\, ds dt$$

$$p$$
$$= P(\text{No Signal})$$
$$= P(X_{(a:m)} \le Y_{(j:n)} \le X_{(b:m)})$$
$$= E_{X_{(a:m)}, X_{(b:m)}} \left(P(X_{(a:m)} \le Y_{(j:n)} \le X_{(b:m)}) \mid X_{(a:m)}, X_{(b:m)} \right)$$
$$= E_{X_{(a:m)}, X_{(b:m)}} \left(P(G(X_{(a:m)}) \le G(Y_{(j:n)}) \le G(X_{(b:m)})) \mid X_{(a:m)}, X_{(b:m)} \right)$$

By the PIT we have that $U_{(j:n)} = G(Y_{(j:n)})$ where G is the continous cdf of the Phase II sample $Y_1, Y_2, ..., Y_n$. Using this we obtain

$$= E_{X_{(a:m)}, X_{(b:m)}} \left(P(G(X_{(a:m)}) \le U_{(j:n)} \le G(X_{(b:m)})) \mid X_{(a:m)}, X_{(b:m)} \right)$$

By the PIT we have that $U_{(a:m)} = F(X_{(a:m)})$ so that $X_{(a:m)} = F^{-1}(U_{(a:m)})$ and $U_{(b:m)} = F(X_{(b:m)})$ so that $X_{(b:m)} = F^{-1}(U_{(b:m)})$ where F is the continous cdf of the reference sample $X_1, X_2, ..., X_m$. Using this we obtain

$$= E_{U_{(a:m)},U_{(b:m)}}\left(P(GF^{-1}(U_{(a:m)}) \le U_{(j:n)} \le GF^{-1}(U_{(b:m)}))|U_{(a:m)},U_{(b:m)}\right)$$

$$= \int_0^1 (GF^{-1}(s) \le U_{(j:n)} \le GF^{-1}(t))|U_{(a:m)} = s, U_{(b:m)} = t) f(s,t) dsdt$$

where $f(s,t)$ is the joint pdf of $U_{(a:m)}$ and $U_{(b:m)}$ which is given by

$$f(s,t) = \frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}$$

$$= \int_0^1 \int_0^t \left(\int_{GF^{-1}(s)}^{GF^{-1}(t)} \frac{1}{\beta(j,n-j+1)} u^{j-1}(1-u)^{n-j} du \right) f(s,t) dsdt$$

$$= \int_0^1 \int_0^t \left(\left(\int_{GF^{-1}(s)}^{GF^{-1}(t)} \frac{1}{\beta(j,n-j+1)} u^{j-1}(1-u)^{n-j} du \right)^k \right) \frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b} dsdt$$

The term $(1-u)^{n-j}$ can be expanded to $(1-u)^{n-j} = \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} 1^{n-j-h} u^h =$ $\sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} u^h$ by using a binomial expansion (see equation 4.15)) and we obtain

$$= \int_0^1 \int_0^t \left(\left(\int_{GF^{-1}(s)}^{GF^{-1}(t)} \frac{1}{\beta(j,n-j+1)} u^{j-1} \left(\sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} u^h \right) du \right)^k \right) \frac{m!}{(a-1)!(b-a-1)!(m-b)!} \times$$
$$s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b} dsdt$$

By taking all the constants out of the integral sign and simplifying by setting $u^{j-1}u^h = u^{j-1+h}$ we obtain

$$= \int_0^1 \int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} (-1)^h \binom{n-j}{h} \int_{GF^{-1}(s)}^{GF^{-1}(t)} u^{j-1+h} du \right)^k \frac{m!}{(a-1)!(b-a-1)!(m-b)!} \times$$
$$s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b} dsdt$$

By integrating $\int_{GF^{-1}(s)}^{GF^{-1}(t)} u^{j-1+h} du = \frac{u^{j+h}}{j+h}\Big|_{u=GF^{-1}(s)}^{u=GF^{-1}(t)} = \frac{\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h}}{j+h}$ we obtain

$$= \int_0^1 \int_0^t \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t) \right)^{j+h} - \left(GF^{-1}(s) \right)^{j+h} \right) \right) \times$$
$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1} (t-s)^{b-a-1} (1-t)^{m-b} \, dsdt .$$

Result 4.27: Probability of a signal - conditional

A signalling event occurs when $Y_{(j:n)} < X_{(a:m)}$ or $Y_{(j:n)} > X_{(b:m)}$.

$$P\left(\text{Signal} \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = 1 - \int_{G(x)}^{G(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1} (1-u)^{n-j} \, du$$

Result 4.28: Probability of a signal - unconditional

$$1 - p = P(\text{Signal}) =$$
$$1 - \int_0^1 \int_0^t \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t) \right)^{j+h} - \left(GF^{-1}(s) \right)^{j+h} \right) \right) \times$$
$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1} (t-s)^{b-a-1} (1-t)^{m-b} \, dsdt$$

Result 4.29: Probability of a false alarm - conditional

$$CFAR = 1 - \int_{F(x)}^{F(z)} \frac{1}{\beta(j, n-j+1)} u^{j-1} (1-u)^{n-j} \, du$$

Follows immediately from Result 4.27, since $G = F$.

Result 4.30: Probability of a false alarm - unconditional

$$FAR = 1 - \int_0^1 \int_0^t \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(t^{j+h} - s^{j+h} \right) \right) \times \frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1} (t-s)^{b-a-1} (1-t)^{m-b} \, ds dt$$

Follows immediately from Result 4.28, since $G = F$ and therefore $GF^{-1}(s) = FF^{-1}(s) = s$ and $GF^{-1}(t) = FF^{-1}(t) = t$.

Result 4.31: Run-length distribution - conditional

$$P\left(N = k \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = (p(x,z))^{k-1}(1 - p(x,z))$$
$$\text{for} \quad k = 1,2,3,...$$

The conditional run length, denoted by $N \mid X_{(a:m)} = x, X_{(b:m)} = z$, will have a geometric distribution with parameter $1 - p(x,z)$, because all the signalling events are independent. Therefore we have that

$N \mid X_{(a:m)} = x, X_{(b:m)} = z \sim GEO(1 - p(x,z))$

$P\left(N = k \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = (p(x,z))^{k-1}(1 - p(x,z))$

for $k = 1,2,3,...$

Consequently, the cumulative distribution function (cdf) is found from

$P\left(N \le k \mid X_{(a:m)} = x, X_{(b:m)} = z\right)$

$= \sum_{i=1}^{k} (p(x,z))^{i-1}(1 - p(x,z)) = 1 - (p(x,z))^k \quad \text{for} \quad k = 1,2,3,...$

We also have that

$P\left(N > k \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = 1 - \left(1 - (p(x,z))^k\right) = (p(x,z))^k$.

Result 4.32: Average run-length - conditional

$$CARL = E\left(N \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = \frac{1}{1 - p(x,z)}$$

or

$$CARL = E\left(N \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = \sum_{k=0}^{\infty} (p(x,z))^k$$

Since the conditional run length, denoted by $N \mid X_{(a:m)} = x, X_{(b:m)} = z$ has a geometric distribution with parameter $1 - p(x,z)$, the conditional average run length is given by

$$CARL = E\left(N \mid X_{(a:m)} = x, X_{(b:m)} = z\right) = \frac{1}{1 - p(x,z)}.$$

The second expression follows immediately from the geometric expansion of $(1 - p(x,z))^{-1}$ for $p(x,z) < 1$.

Result 4.33: Run-length distribution - unconditional

$$P(N = k) = D^*(k-1) - D^*(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D^*(0) = 1$$

and

$$D^*(k) = \int_0^1 \int_0^t \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right) \right)^k \times$$

$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1} (t-s)^{b-a-1} (1-t)^{m-b} \, ds\, dt$$

$$\begin{aligned}
&P(N = k) \\
&= E_{X_{(a:m)}, X_{(b:m)}} \left(P\left(N = k \mid X_{(a:m)} = x, X_{(b:m)} = z\right) \right) \\
&= E_{X_{(a:m)}, X_{(b:m)}} \left((p(x,z))^{k-1} (1 - p(x,z)) \right) \\
&= E_{X_{(a:m)}, X_{(b:m)}} \left((p(x,z))^{k-1} - (p(x,z))^k \right) \\
&= E_{X_{(a:m)}, X_{(b:m)}} \left((p(x,z))^{k-1} \right) - E_{X_{(a:m)}, X_{(b:m)}} \left((p(x,z))^k \right)
\end{aligned}$$

By only focussing on $E_{X_{(a:m)}, X_{(b:m)}} \left((p(x,z))^k \right)$ we have

$$E_{X_{(a:m)},X_{(b:m)}}\left((p(x,z))^k\right)$$

$$=E_{U_{(a:m)},U_{(b:m)}}\left(\left(\int_{GF^{-1}(s)}^{GF^{-1}(t)}\frac{1}{\beta(j,n-j+1)}u^{j-1}(1-u)^{n-j}\,du\right)^k\right)$$

$$=\int_0^1\int_0^t\left(\left(\int_{GF^{-1}(s)}^{GF^{-1}(t)}\frac{1}{\beta(j,n-j+1)}u^{j-1}(1-u)^{n-j}\,du\right)^k\right)f(s,t)\,dsdt$$

where $f(s,t)$ is the joint pdf of $U_{(a:m)}$ and $U_{(b:m)}$ which is given by

$$f(s,t)=\frac{m!}{(a-1)!(b-a-1)!(m-b)!}s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}$$

$$=\int_0^1\int_0^t\left(\left(\int_{GF^{-1}(s)}^{GF^{-1}(t)}\frac{1}{\beta(j,n-j+1)}u^{j-1}(1-u)^{n-j}\,du\right)^k\right)\frac{m!}{(a-1)!(b-a-1)!(m-b)!}s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

The term $(1-u)^{n-j}$ can be expanded to $(1-u)^{n-j}=\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}1^{n-j-h}u^h=$ $\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}u^h$ by using a binomial expansion (see equation (4.15)) and we obtain

$$=\int_0^1\int_0^t\left(\left(\int_{GF^{-1}(s)}^{GF^{-1}(t)}\frac{1}{\beta(j,n-j+1)}u^{j-1}\left(\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}u^h\right)du\right)^k\right)\frac{m!}{(a-1)!(b-a-1)!(m-b)!}\times$$
$$s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

By taking all the constants out of the integral sign and simplifying by setting $u^{j-1}u^h=u^{j-1+h}$ we obtain

$$=\int_0^1\int_0^t\left(\frac{1}{\beta(j,n-j+1)}\sum_{h=0}^{n-j}(-1)^h\binom{n-j}{h}\int_{GF^{-1}(s)}^{GF^{-1}(t)}u^{j-1+h}\,du\right)^k\frac{m!}{(a-1)!(b-a-1)!(m-b)!}\times$$
$$s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

By integrating $\int_{GF^{-1}(s)}^{GF^{-1}(t)} u^{j-1+h}du = \frac{u^{j+h}}{j+h}\Big|_{u=GF^{-1}(s)}^{u=GF^{-1}(t)} = \frac{\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h}}{j+h}$ we obtain

$$= \int_0^1\int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right) \right)^k \times$$
$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

$$= D^*(k),$$

say.

Recall that $P(N=k) = E_{X_{(a:m)},X_{(b:m)}}\left((p(x,z))^{k-1}\right) - E_{X_{(a:m)},X_{(b:m)}}\left((p(x,z))^{k}\right)$

Therefore, $P(N=k) = D^*(k-1) - D^*(k)$ for $k = 1,2,3,\ldots,$ and $D^*(0) = 1.$

$D^*(0) = 1$ since

$$D^*(0) = \int_0^1\int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right) \right)^0 \times$$
$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

$$= \int_0^1\int_0^t \frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$
$$= \int_0^1\int_0^t f(s,t)\,dsdt$$
$$= 1.$$

Result 4.34: In-control run-length distribution

$$P_C(N=k) = D(k-1) - D(k) \quad \text{for} \quad k = 1,2,3,\ldots, \quad D(0) = 1$$

and

$$D(k) = \int_0^1\int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(t^{j+h} - s^{j+h} \right) \right)^k \frac{m!}{(a-1)!(b-a-1)!(m-b)!} \times$$
$$s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

Recall that the reference sample of size m, $X_1, X_2, ..., X_m$, is available from an in-control process with a continuous cdf, $F(x)$. The plotting statistic $Y^h_{(j:n)}$ is the j^{th} order statistic from the h^{th} Phase II sample of size n_h. Let $G^h(y)$ denote the cdf of the distribution of the h^{th} Phase II sample. A process is said to be in-control at stage h when $G^h = F$. Assume that the Phase II samples are all of the same size, n, so that the subscript h can be suppressed. Therefore, a process is said to be in-control when $G = F$. Therefore, the in-control run length distribution is obtained by setting $G = F$ into the equation for the out-of-control run length distribution.

The out-of-control run length distribution for the two-sided chart is given by

$$P(N = k) = D^*(k-1) - D^*(k) \quad \text{for} \quad k = 1,2,3,..., \quad D^*(0) = 1$$

and (4.20)

$$D^*(k) = \int_0^1 \int_0^t \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right) \right)^k \times$$

$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\, ds dt .$$

Therefore, the in-control run length distribution for the two-sided chart is obtained by setting $G = F$ into equation (4.20) and we obtain

$$P_C(N = k) = D(k-1) - D(k) \quad \text{for} \quad k = 1,2,3,..., \quad D(0) = 1$$

and (4.21)

$$D(k) = \int_0^1 \int_0^t \left(\frac{1}{\beta(j, n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(t^{j+h} - s^{j+h}\right) \right)^k \frac{m!}{(a-1)!(b-a-1)!(m-b)!} \times$$

$$s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\, ds dt .$$

Result 4.35: Out-of-control average run-length - unconditional

$$UARL_{\delta} = \int_0^1 \int_0^t \frac{1}{1 - S(s,t,j,n,F,G)} f(s,t)dsdt$$

with

$$S(s,t,j,n,F,G) = \frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right)$$

Let $UARL_{\delta}$ denote the unconditional average run length, where δ refers to the out-of-control case. To derive an expression for the $UARL_{\delta}$, recall that

$$E_{X_{(a:m)}, X_{(b:m)}} \left((p(x,z))^k \right)$$
$$= D^*(k)$$
$$= \int_0^1 \int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right) \right)^k \times$$
$$\frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1} (t-s)^{b-a-1} (1-t)^{m-b} dsdt .$$

$$= \int_0^1 \int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right) \right)^k f(s,t)dsdt$$

Let $S(s,t,j,n,F,G) = \frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(\left(GF^{-1}(t)\right)^{j+h} - \left(GF^{-1}(s)\right)^{j+h} \right)$

$$= \int_0^1 \int_0^t \left(S(s,t,j,n,F,G) \right)^k f(s,t)dsdt .$$

Finally, we have that (from the second expression in Result 4.32)

$$UARL_{\delta}$$
$$= \sum_{k=0}^{\infty} E_{X_{(a:m)}, X_{(b:m)}} \left((p(x,z))^k \right)$$
$$= \sum_{k=0}^{\infty} D^*(k)$$

$$= \sum_{k=0}^{\infty} \int_0^1 \int_0^t \left(S(s,t,j,n,F,G)\right)^k f(s,t)dsdt$$

$$= \int_0^1 \int_0^t \sum_{k=0}^{\infty} \left(S(s,t,j,n,F,G)\right)^k f(s,t)dsdt$$

$$= \int_0^1 \int_0^t \frac{1}{1-S(s,t,j,n,F,G)} f(s,t)dsdt\,.$$

Result 4.36: In-control average run-length - unconditional

$$UARL_0 = \int_0^1 \int_0^t \left(\frac{1}{1-S(s,t,j,n)}\right) \frac{m!}{(a-1)!(b-a-1)!(m-b)!} s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

with

$$S(s,t,j,n) = \frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(t^{j+h} - s^{j+h}\right)$$

Let $UARL_0$ denote the unconditional average run length, where 0 refers to the in-control case. To derive an expression for the $UARL_0$, recall that the in-control run length distribution for the two-sided chart is given by

$P_C(N=k) = D(k-1) - D(k)$ for $k = 1,2,3,\ldots,$ $D(0) = 1$

with

$$D(k) = \int_0^1 \int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(t^{j+h} - s^{j+h}\right)\right)^k \frac{m!}{(a-1)!(b-a-1)!(m-b)!} \times$$

$$s^{a-1}(t-s)^{b-a-1}(1-t)^{m-b}\,dsdt$$

$$= \int_0^1 \int_0^t \left(\frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(t^{j+h} - s^{j+h}\right)\right)^k f(s,t)dsdt$$

For simplicity let $S(s,t,j,n) = \frac{1}{\beta(j,n-j+1)} \sum_{h=0}^{n-j} \frac{(-1)^h}{j+h} \binom{n-j}{h} \left(t^{j+h} - s^{j+h}\right)$, therefore we obtain

$$D(k) = \int_0^1 \int_0^t \left(S(s,t,j,n)\right)^k f(s,t)dsdt$$

Finally, we have that

$$
\begin{aligned}
&UARL_0 \\
&= \sum_{k=0}^{\infty} D(k) \\
&= \sum_{k=0}^{\infty} \int_0^1 \int_0^t \left(S(s,t,j,n)\right)^k f(s,t)dsdt \\
&= \int_0^1 \int_0^t \sum_{k=0}^{\infty} \left(S(s,t,j,n)\right)^k f(s,t)dsdt \\
&= \int_0^1 \int_0^t \frac{1}{1-S(s,t,j,n)} f(s,t)dsdt .
\end{aligned}
$$

4.1.10. Run-length distribution and *ARL* under some alternatives

In the nonparametric setting, we consider, more generally, monitoring the center value or the location parameter and/or a scale parameter of a process. The location parameter represents a typical value and could be the mean or the median or some other percentile of the distribution; the latter two are especially attractive when the underlying distribution is skewed. When the underlying distribution is symmetric, the mean and the median are the same. Also in the nonparametric setting, the processes are implicitly assumed to follow (i) a location model, with a cdf $F(x-\theta)$, where θ is the location parameter or (ii) a scale model, with a cdf $F\left(\frac{x}{\tau}\right)$, where $\tau(>0)$ is the scale parameter. Even more generally, one might consider (iii) the location-scale model with cdf $F\left(\frac{x-\theta}{\tau}\right)$, where θ and τ are the location and the scale parameter, respectively.

Recall that the reference sample is available from an in-control process with a continuous cdf, $F(x)$, and that $G(y)$ denotes the cdf of the distribution of the Phase II sample. The run length distribution depends on F and G, through the function $\psi = GF^{-1}$. A process is said to be in-control when $G = F$. In this case $\psi(u) = G\left(F^{-1}(u)\right) = F\left(F^{-1}(u)\right) = u$.

4.1.10.1. Location alternatives

$F(x) = H(x-\theta_1)$ and $G(x) = H(x-\theta_2)$, where H is a continuous cdf, $x \in \Re$ and $\theta_1, \theta_2 \in \Re$, $\psi(u) = H(\theta_1 - \theta_2 + H^{-1}(u))$. For example, let both F and G be normally distributed with a change in the mean, i.e. $F(x) = \Phi(x)$ and $G(x) = \Phi(x-\theta)$. But $\psi(u) = G(F^{-1}(u))$ (by definition) and therefore $\psi(u) = \Phi(\Phi^{-1}(u) - \theta)$.

4.1.10.2. Scale alternatives

$F(x) = H\left(\frac{x}{\gamma_1}\right)$ and $G(x) = H\left(\frac{x}{\gamma_2}\right)$, where H is a continuous cdf, $x \in \Re$ and $\gamma_1, \gamma_2 \in \Re^+$, $\psi(u) = H\left(\frac{\gamma_1}{\gamma_2} H^{-1}(u)\right)$. For example, let both F and G be normally distributed with a change in the spread, i.e. $F(x) = \Phi(x)$ and $G(x) = \Phi\left(\frac{x}{\gamma}\right)$. But $\psi(u) = G(F^{-1}(u))$ (by definition) and therefore $\psi(u) = \Phi\left(\frac{\Phi^{-1}(u)}{\gamma}\right)$.

4.1.10.3. Location-scale alternatives

$F(x) = H\left(\frac{x-\theta_1}{\gamma_1}\right)$ and $G(x) = H\left(\frac{x-\theta_2}{\gamma_2}\right)$, where H is a continuous cdf, $x \in \Re$, $\theta_1, \theta_2 \in \Re$, and $\gamma_1, \gamma_2 \in \Re^+$, $\psi(u) = H\left(\frac{\theta_1 - \theta_2}{\gamma_2} + \frac{\gamma_1}{\gamma_2} H^{-1}(u)\right)$. For example, let both F and G be normally distributed with a change in the mean and spread, i.e. $F(x) = \Phi(x)$ and $G(x) = \Phi\left(\frac{x-\theta}{\gamma}\right)$. But $\psi(u) = G(F^{-1}(u))$ (by definition) and therefore $\psi(u) = \Phi\left(\frac{\Phi^{-1}(u) - \theta}{\gamma}\right) = \Phi\left(\frac{\Phi^{-1}(u)}{\gamma} - \frac{\theta}{\gamma}\right)$.

4.1.10.4. Lehmann alternatives

$G(x) = F^{\delta}(x)$, where $x \in \Re$ and $\delta \in \Re^{+}$, $\psi(u) = u^{\delta}$. For example, let $F(x) = u$ and $G(x) = (F(x))^{\delta}$. But $\psi(u) = G(F^{-1}(u))$ (by definition) and therefore $\psi(u) = (F(F^{-1}(u)))^{\delta} = u^{\delta}$.

- For $\delta = 1$: $\psi(u) = u = F(x)$.
- For $\delta = 2$: $\psi(u) = u^{2} = F^{2}(x)$.

4.1.10.5. Proportional hazards alternatives

$G(x) = 1 - (1 - F(x))^{\gamma}$, where $x \in \Re$ and $\gamma \in \Re^{+}$, $\psi(u) = 1 - (1 - u)^{\gamma}$. For example, let $F(x) = u$ and $G(x) = 1 - (1 - F(x))^{\gamma}$. But $\psi(u) = G(F^{-1}(u))$ (by definition) and therefore $\psi(u) = 1 - (1 - F(F^{-1}(u)))^{\gamma} = 1 - (1 - u)^{\gamma}$.

4.1.10.6. Summary

Although a lot of research has been done in the last few years regarding Lehmann and proportional hazard alternatives (see for example Van der Laan and Chakraborti (1999)), more remains to be done. Van der Laan and Chakraborti (1999) showed that the power of a precedence test can be determined for both the Lehmann and proportional hazards alternatives. The body of literature on Lehmann and proportional hazards alternatives is growing. However, in our opinion, a discussion on this topic is better postponed for the future.

4.2. The Shewhart-type control chart with runs-type signalling rules

4.2.1. Introduction

Chakraborti, Eryilmaz and Human (2006) considered enhancing the precedence charts with *2-of-2* type signalling rules. The *2-of-2* DR and *2-of-2* KL rules were defined previously (see Section 3.2). Recall that the *2-of-2* KL chart signals when two of the most recent charting

statistics both fall either on or above or on or below the control limits, whereas the *2-of-2* DR chart signals when the charting statistics fall either both on or above or both on or below or one on or above (below) and the next one on or below (above) the control limits. We illustrate these procedures using the Montgomery (2001) piston ring data.

4.2.2. Example

Example 4.1

A sign-like control chart based on the Montgomery (2001) piston ring data

We illustrate the sign-like control charts using a set of data from Montgomery (2001, Tables 5.1 and 5.2) on inside diameters of piston rings manufactured by a forging process. Table 5.1 of Montgomery (2001) contains 25 retrospective or Phase I samples, each of size five, that were collected when the process was thought to be in-control. When working with individual observations, we have $25\times 5 = 125$, i.e. $m = 125$, individual observations. Table 5.2 of Montgomery (2001) contains 15 prospective or Phase II samples, each of five observations.

In order to implement the control charts, the charting constants are needed. Generally, one finds the chart constants so that a specified ARL_0, such as 500 or 370, is obtained. For the precedence type charts, symmetric control limits are used so that $b = m - a + 1$ and only one charting constant a (≥ 1) needs to be found. Possible control limits for the three charts are shown in Table 4.2 for $m = 125$, $n = 5$ and $j = 3$, along with the corresponding *FAR* and ARL_0 values. The basic Shewhart-type precedence chart is referred to as the *1-of-1* chart.

Table 4.2. In-control average run length (ARL_0), false alarm rate (*FAR*) and chart constant (a) for the *1-of-1*, *2-of-2* DR and *2-of-2* KL precedence charts when $m = 125$, $n = 5$ and $j = 3$ [*].

	1-of-1		*2-of-2* DR			*2-of-2* KL		
a	ARL_0	***FAR***	***a***	ARL_0	***FAR***	***a***	ARL_0	***FAR***
5	1315.98	0.001865	19	464.38	0.004020	19	819.47	0.002355
6	695.09	0.002948	20	344.73	0.005195	20	608.81	0.003019
7	413.80	0.004358	21	260.69	0.006627	21	460.54	0.003823
8	267.40	0.006164	22	200.46	0.008356	22	354.09	0.004788

Thus, for an ARL_0 of 500, one can take $a = 7$ and $b = 119$ so that the control limits for the *1-of-1* precedence chart are the 7^{th} and the 119^{th} ordered values of the reference sample. Thus $L\hat{C}L = X_{(7:125)} = 73.984$ and $U\hat{C}L = X_{(119:125)} = 74.017$, which yield an in-control average run length of 413.80 and a *FAR* of 0.0044. A plot of the sample medians for the *1-of-1* chart is shown in Figure 4.3. It is seen that the *1-of-1* precedence chart signals on the 12^{th} sample in the prospective phase.

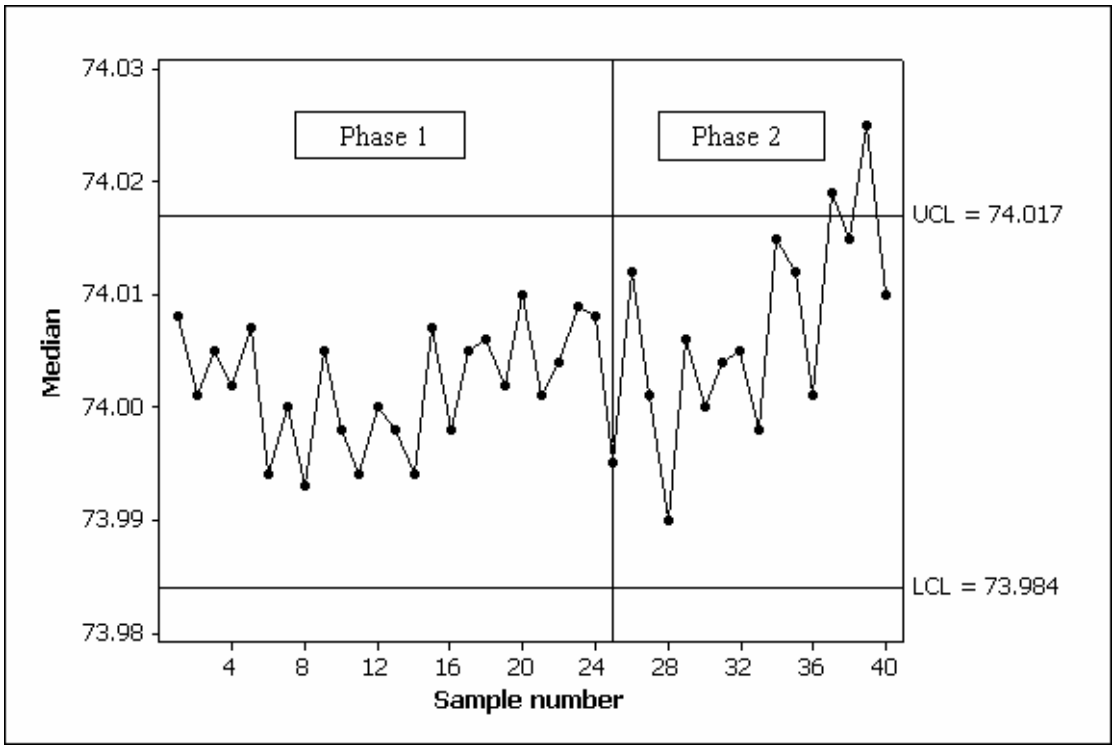

Figure 4.3. *1-of-1* Precedence chart for the Montgomery (2001) piston ring data.

[*] Table 4.2 appears in Chakraborti, Eryilmaz and Human (2006), Table 3.

For the *2-of-2* DR chart, take $a = 19$ so that $b = 125 - 19 + 1 = 107$ and the resulting limits, $L\hat{C}L = X_{(19:125)} = 73.992$ and $U\hat{C}L = X_{(107:125)} = 74.012$, yield an ARL_0 and *FAR* of 464.38 and 0.0040, respectively. Note however that if one chooses $a = 20$ so that $b = 106$, the control limits are $L\hat{C}L = X_{(20:125)}$ and $U\hat{C}L = X_{(106:125)}$ and the ARL_0 decreases to 344.73, whereas the *FAR* slightly increases to 0.0052. The *2-of-2* DR chart is shown in Figure 4.4.

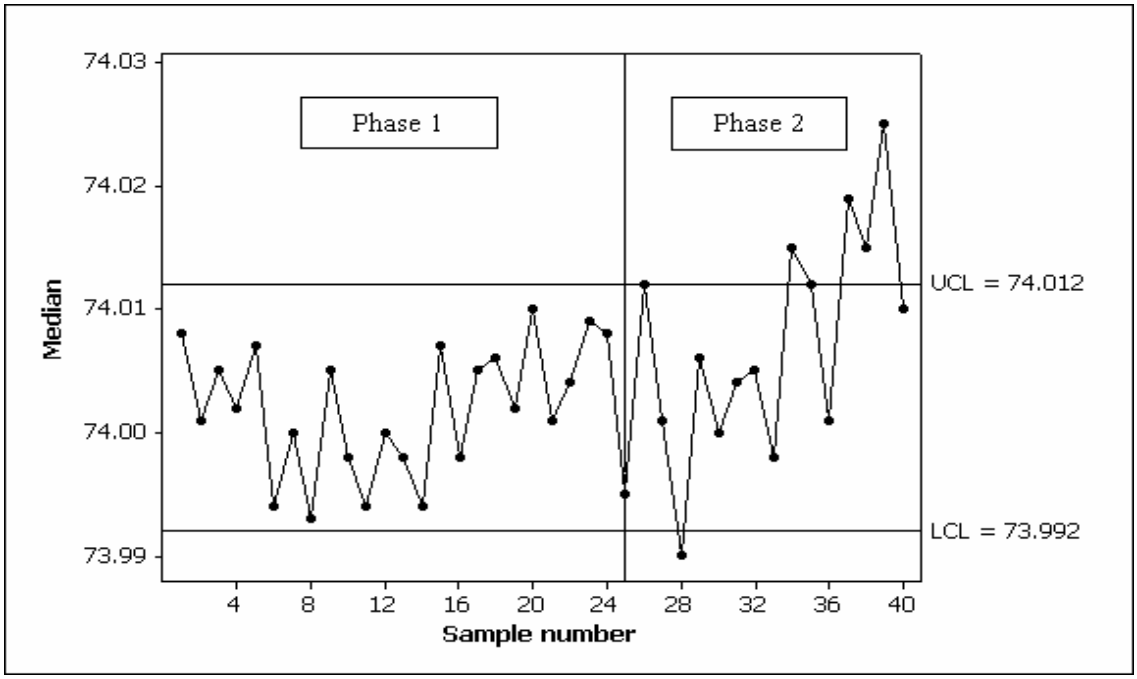

Figure 4.4. 2-*of-2* DR precedence chart for the Montgomery (2001) piston ring data.

For the *2-of-2* KL chart take $a = 21$ so that $b = 125 - 21 + 1 = 105$ so that $L\hat{C}L = X_{(21:125)} = 73.992$ and $U\hat{C}L = X_{(105:125)} = 74.011$, and this yields an ARL_0 of 460.54 and a *FAR* of 0.0038, respectively. This *2-of-2* KL chart is almost identical to the DR chart in Figure 4.4.

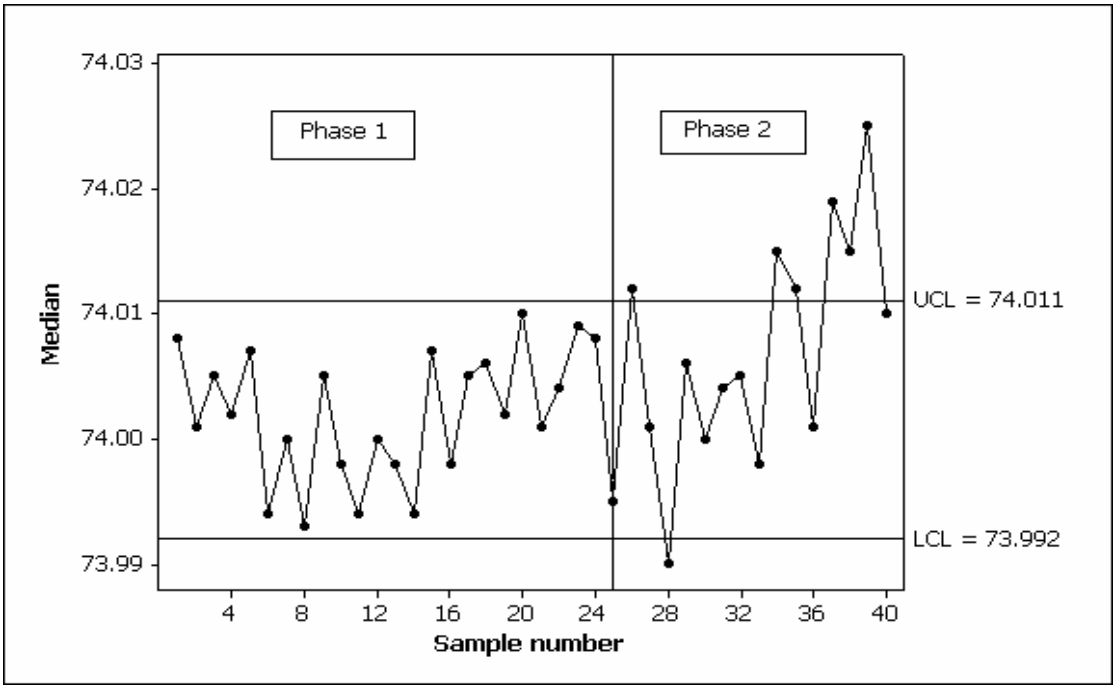

Figure 4.5. 2-*of*-2 KL precedence chart for the Montgomery (2001) piston ring data.

Both the *2-of-2* DR and KL charts signal on the 10th sample in the prospective phase. Note, however, that the achieved *FAR* values for all three charts are much larger than the nominal *FAR* of 0.0027.

4.2.3. Summary

In this chapter we examined sign-like control charts with runs-type signalling rules. We illustrated these procedures using the piston ring data from Montgomery (2001) to help the reader to understand the subject more thoroughly. There are many advantages to using these nonparametric charts (see Section 1.4). Chakraborti, Eryilmaz and Human (2006) draw attention to two advantages in particular, namely, that these charts can be applied as soon as the required order statistics are observed (recall that both the control limits and the charting statistic are based on order statistics), whereas for the Shewhart $\overline{X}$ charts one needs the full dataset to calculate the average. Moreover, these charts can be adapted to and applied in the case of ordinal data. As a result Chakraborti, Eryilmaz and Human (2006) recommend that these charts be used in practice.

Chapter 5: Signed-rank-like charts

5.1. The Shewhart-type control chart

5.1.1. Introduction

The statistics used in nonparametric control charts are mostly signs, ranks and signed-ranks and related to nonparametric procedures, such as the Wilcoxon signed-rank test and the Mann-Whitney-Wilcoxon rank-sum test. When considering nonparametric tests based on ranks, such tests deal with the ranking of independent, identically distributed (iid) random variables (under the assumption that the process is in-control). In Chapter 5 we consider nonparametric tests that involve ranking random variables that are exchangeable (again, this holds under the assumption that the process is in-control), meaning that each possible ranking is equally likely. Randles and Wolfe (1979) state that the term *rank-like* is used to describe a type of test procedure where the variables that are ranked are not the original observations, but are, instead, functions of them. The term *rank-like* was first introduced by Moses (1963). Moses's rank-like test is a nonparametric test for comparing differences in dispersion between two samples in which the medians are not equal. This requires randomly allocating the sample observations into two subgroups, ranking the subgroups according to their dispersion indexes and calculating the ranks sums for each subgroup. It should be noted that although Moses's rank-like test uses rankings of iid random variables (under the assumption that the process is in-control), these variables are not the original observations, but instead, functions of them. Bakir (2006) considered what are called *signed-rank-like* (SRL) statistics and used these to construct distribution-free charts. He uses the median of a reference sample (taken when the process was operating in-control) to estimate the unknown in-control process center.

5.1.2. Definition of the signed-rank-like test statistic

Assume that a reference sample of size $m > 1$, $X_1, X_2, ..., X_m$, is available from an in-control process with an unknown continuous cdf $F(x)$. Let $Y_{i1}, Y_{i2}, ..., Y_{in}$, $i = 1,2,...,$ denote the i^{th} test sample of size n. In case U the median of the in-control distribution (assumed to be symmetric) is unknown and can be estimated by the median of a reference sample, say M.

Let R_{ij}^{*} denote the rank of $\left|y_{ij}-M\right|$ within the subgroup $\left(\left|y_{i1}-M\right|,...,\left|y_{in}-M\right|\right)$ for $i=1,2,3...$. R_{ij}^{*} can be calculated using

$$R_{ij}^{*}=1+\sum_{k=1}^{n}(I\mid y_{ik}-M\mid<\mid y_{ij}-M\mid)\ \text{for}\ j=1,2,...,n \tag{5.1}$$

where I is the indicator function defined by $I(x)=1, 0$ if x is true or false.

The charting statistic is given by

$$SRL_i=\sum_{j=1}^{n}sign\left(y_{ij}-M\right)R_{ij}^{*}\ \text{ for }\ i=1,2,3... \tag{5.2}$$

where $sign(x)$ = -1, 0, 1 if $x<0$, $=0$, >0. The charting statistic, SRL_i, is a direct analog of the plotting statistic SR_i used in case K. If the charting statistic SRL_i falls between the two control limits, that is, $LCL<SRL_i<UCL$, the process is considered to be in-control. If the charting statistic SRL_i falls on or outside one of the control limits, that is $SRL_i\le LCL$ or $SRL_i\ge UCL$, the process is considered to be out-of-control.

Example 5.1

A Shewhart-type signed-rank-like statistic for the Montgomery (2001) piston ring data

We illustrate the Shewhart-type signed-rank-like chart using a set of data from Montgomery (2001; Tables 5.1 and 5.2) on the inside diameters of piston rings manufactured by a forging process. Table 5.1 of Montgomery (2001) contains the reference sample of size $m=125$ (see example 4.1 for an explanation of why m is equal to 125 (and not 25 like some of the earlier examples) and the median of this reference sample equals 74.001, i.e. $M=74.001$.

Panel *a* of Table 5.1 exhibits the individual observations of 15 independent samples, each of size 5 i.e. $n=5$. The absolute deviations $\left|y_{ij}-M\right|$ and $sign\left(y_{ij}-M\right)$ are shown in panel *b* and panel *c* of Table 5.1, respectively. The rank R_{ij}^{*} and the $sign\left(y_{ij}-M\right)R_{ij}^{*}$ values are shown in panel *a* and panel *b* of Table 5.2, respectively. Panel *c* of Table 5.2 holds the SRL-values i.e. SRL_i for $i=1,2,3,...,15$.

Table 5.1. Data and calculations for the signed-rank-like chart.*

Sample number	Panel *a*					Panel *b*					Panel *c*				
	y_{i1}	y_{i2}	y_{i3}	y_{i4}	y_{i5}	$\lvert y_{i1} - M \rvert$	$\lvert y_{i2} - M \rvert$	$\lvert y_{i3} - M \rvert$	$\lvert y_{i4} - M \rvert$	$\lvert y_{i5} - M \rvert$	$sign(y_{i1} - M)$	$sign(y_{i2} - M)$	$sign(y_{i3} - M)$	$sign(y_{i4} - M)$	$sign(y_{i5} - M)$
1	74.012	74.015	74.030	73.986	74.000	0.011	0.014	0.029	0.015	0.001	1	1	1	-1	-1
2	73.995	74.010	73.990	74.015	74.001	0.006	0.009	0.011	0.014	0.000	-1	1	-1	1	0
3	73.987	73.999	73.985	74.000	73.990	0.014	0.002	0.016	0.001	0.011	-1	-1	-1	-1	-1
4	74.008	74.010	74.003	73.991	74.006	0.007	0.009	0.002	0.01	0.005	1	1	1	-1	1
5	74.003	74.000	74.001	73.986	73.997	0.002	0.001	0.000	0.015	0.004	1	-1	0	-1	-1
6	73.994	74.003	74.015	74.020	74.004	0.007	0.002	0.014	0.019	0.003	-1	1	1	1	1
7	74.008	74.002	74.018	73.995	74.005	0.007	0.001	0.017	0.006	0.004	1	1	1	-1	1
8	74.001	74.004	73.990	73.996	73.998	0.000	0.003	0.011	0.005	0.003	0	1	-1	-1	-1
9	74.015	74.000	74.016	74.025	74.000	0.014	0.001	0.015	0.024	0.001	1	-1	1	1	-1
10	74.030	74.005	74.000	74.016	74.012	0.029	0.004	0.001	0.015	0.011	1	1	-1	1	1
11	74.001	73.990	73.995	74.010	74.024	0.000	0.011	0.006	0.009	0.023	0	-1	-1	1	1
12	74.015	74.020	74.024	74.005	74.019	0.014	0.019	0.023	0.004	0.018	1	1	1	1	1
13	74.035	74.010	74.012	74.015	74.026	0.034	0.009	0.011	0.014	0.025	1	1	1	1	1
14	74.017	74.013	74.036	74.025	74.026	0.016	0.012	0.035	0.024	0.025	1	1	1	1	1
15	74.010	74.005	74.029	74.000	74.020	0.009	0.004	0.028	0.001	0.019	1	1	1	-1	1

* See SAS Program 10 in Appendix B for the calculation of the values in Table 5.1.

Table 5.2. Calculations for the signed-rank-like chart*.

Sample number	Panel *a*					Panel *b*					Panel *c*
	R_{i1}^*	R_{i2}^*	R_{i3}^*	R_{i4}^*	R_{i5}^*	$sign(y_{i1}-M)R_{i1}^*$	$sign(y_{i2}-M)R_{i2}^*$	$sign(y_{i3}-M)R_{i3}^*$	$sign(y_{i4}-M)R_{i4}^*$	$sign(y_{i5}-M)R_{i5}^*$	SRL_i
1	2	3	5	4	1	2	3	5	-4	-1	5
2	2	3	4	5	1	-2	3	-4	5	0	2
3	4	2	5	1	3	-4	-2	-5	-1	-3	-15
4	3	4	1	5	2	3	4	1	-5	2	5
5	3	2	1	5	4	3	-2	0	-5	-4	-8
6	3	1	4	5	2	-3	1	4	5	2	9
7	4	1	5	3	2	4	1	5	-3	2	9
8	1	2.5	5	4	2.5	0	2.5	-5	-4	-2.5	-9
9	3	1.5	4	5	1.5	3	-1.5	4	5	-1.5	9
10	5	2	1	4	3	5	2	-1	4	3	13
11	1	4	2	3	5	0	-4	-2	3	5	2
12	2	4	5	1	3	2	4	5	1	3	15
13	5	1	2	3	4	5	1	2	3	4	15
14	2	1	5	3	4	2	1	5	3	4	15
15	3	2	5	1	4	3	2	5	-1	4	13

The control limits are chosen to give a certain false alarm rate or in-control ARL. A symmetric two-sided chart is obtained by choosing $LCL = -UCL$. For $n = 5$, the control limits for the signed-rank-like chart are set at ± 15. These control limits yield an in-control ARL of 16 and a FAR of 0.0626 (these values were obtained by the use of a simulation study (see SAS Program 6 in Appendix B) where $m = 500$ and $n = 5$). With such a small in-control average run length, many false alarms will be signalled by this chart leading to possible loss of time and resources. The chart is shown in Figure 5.1 with control limits at ± 15.

* See SAS Program 10 Appendix B for the calculation of the values in Table 5.2.

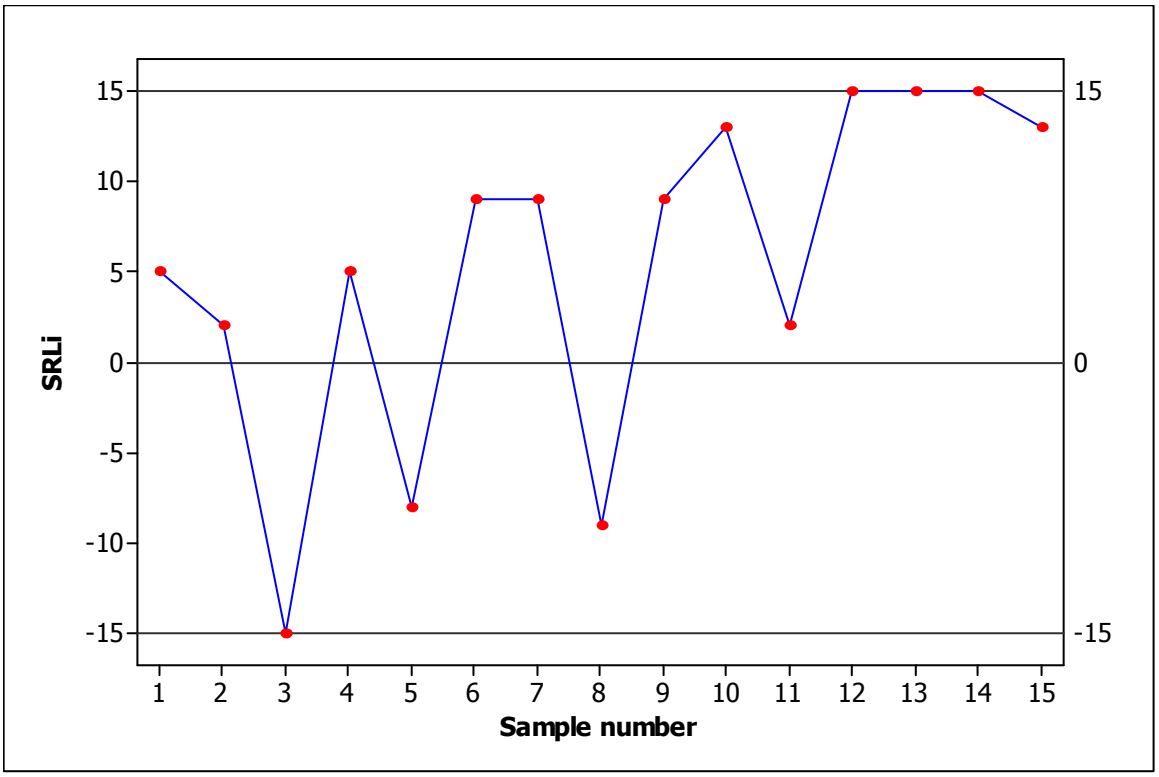

Figure 5.1. Shewhart-type signed-rank-like control chart for Montgomery (2001) piton ring data.

Observations 3, 12, 13 and 14 lie on the upper control limit which indicates that the process is out-of-control starting at sample 3. It appears most likely that the process median has shifted upwards from the target value of 74*mm*. Corrective action and a search for assignable causes is necessary.

5.1.3. Distribution-free properties

We want to establish that the charting statistic SRL_i is distribution-free. If the latter is true, then the signed-rank-like chart based on the SRL_i statistic will be distribution-free. To establish that SRL_i is distribution-free, we first have to look at some properties. Randles and Wolfe (1979) provided various definitions and theorems that are useful in this text.

Definition 1

(See Definition 1.3.1. of Randles and Wolfe (1979), pg. 13)

Two random variables S and T are said to be *equal in distribution* if they have the same cdf. To denote 'equal in distribution' we use the notation $S \overset{d}{=} T$.

Definition 2

(See Definition 1.3.6 of Randles and Wolfe (1979), pg. 15)

A collection of random variables $X_1, X_2, ..., X_n$ is said to be *exchangeable* if for every permutation $(\alpha_1, \alpha_2, ..., \alpha_n)$ of the integers $(1,2,...,n)$, $(X_1, X_2, ..., X_n) \stackrel{d}{=} (X_{\alpha_1}, X_{\alpha_2}, ..., X_{\alpha_n})$.

Theorem 1

(See Theorem 1.3.7 of Randles and Wolfe (1979), pg. 16)

If $\underline{X} \stackrel{d}{=} \underline{Y}$ and $U(\cdot)$ is a (measurable) function (possibly vector valued) defined on the common support of these random variables, then $U(\underline{X}) \stackrel{d}{=} U(\underline{Y})$.

Theorem 2

(See Theorem 11.2.3 of Randles and Wolfe (1979), pg. 356)

Let $\underline{X}_i = (X_{i1}, X_{i2}, ..., X_{ip})$, $i = 1,2,...,n$ be a random sample from some p-variate continuous distribution. Let $g(\cdot)$ be any function of n p-vectors that is symmetric in its arguments. Let $h(\cdot,\cdot)$ be any real-valued function of a p-tuple and the function values of $g(\cdot)$ and define the random variables $W_i = h(\underline{X}_i, g(\underline{X}_1, ..., \underline{X}_n))$, $i = 1,2,...,n$. Then $W_1, W_2, ..., W_n$ are exchangeable random variables, i.e. $(W_1, W_2, ..., W_n) \stackrel{d}{=} (W_{\alpha_1}, W_{\alpha_2}, ..., W_{\alpha_n})$ where $(\alpha_1, \alpha_2, ..., \alpha_n)$ is any permutation of $(1,2,...,n)$.

Theorem 2 can be generalized to complement our problem. Suppose $\underline{X}_1 = \underline{X} = (X_1, X_2, ..., X_m) \sim F$ and $\underline{X}_2 = \underline{Y} = (Y_1, Y_2, ..., Y_n) \sim G$ are independent random samples and F and G are continuous distributions. Let $g(\cdot)$ be a function of $\underline{X}$ that is symmetric in its arguments and let $h(\cdot,\cdot)$ be any real-valued function of $\underline{Y}$ and the function values of $g(\cdot)$. Then define $W_j = h(Y_j, g(\underline{X})) = h(Y_j, g(X_1, ..., X_m))$ for $j = 1, ..., n$. Then, from Theorem 2, we have that $W_1, W_2, ..., W_n$ are exchangeable random variables when $F = G$.

Corollary 1

(See Corollary 2.4.5 of Randles and Wolfe (1979), pg. 50)

Let $S(\underline{\Psi}, \underline{R}^*)$ be a statistic that depends on the observations $X_1, X_2, ..., X_n$ only through $\Psi_1, \Psi_2, ..., \Psi_n$ and $\underline{R}^*$. Then the statistic $S(\cdot)$ is distribution-free over Θ, the collection of joint distributions of n iid continuous random variables, each symmetrically distributed about zero.

Corollary 2

(See Corollary 11.2.5 of Randles and Wolfe (1979), pg. 357)

Let $W_1, W_2, ..., W_n$ be defined as in Theorem 2 and let R_i^* denote the rank of W_i among $W_1, W_2, ..., W_n$. If $P(W_i = W_j) = 0$ for every $i \neq j$, then $P(\underline{R}^* = \underline{r}) = \frac{1}{n!}$ for every $\underline{r}$, a permutation of the integers $(1, ..., n)$. Thus any statistic that is a function of the sample observations $\underline{X}_1, ..., \underline{X}_n$ only through the ranks $R_i^*, ..., R_n^*$ is nonparametric distribution-free over the class of all p-variate continuous distribution.

Lemma 1

(See Lemma 2.4.2 of Randles and Wolfe (1979), pg. 49)

Let Z be a continuous random variable with a distribution that is symmetric about 0. Then the random variables $|Z|$ and $\Psi = \Psi(Z)$ are stochastically independent.

Establishing that the charting statistic SRL_i is distribution-free for an in-control process

The first step in establishing that the charting statistic SRL_i is distribution-free, is by proving that when the process is in-control, i.e. $F = G$, $V_1, V_2, ..., V_n$ are exchangeable random variables, where V_j is defined as $V_j = |Y_j - M|$, $j = 1,2,...,n$, and M is the median of $X_1, ..., X_m$. The proof to this follows from Theorem 2 by setting $g(\underline{X}) = g(X_1, ..., X_m) = M$ and $W_j = h(Y_j, g(X_1, ..., X_m)) = |Y_j - M|$.

The second step in establishing that the charting statistic SRL_i is distribution-free, is by proving that when $F = G$, $U_1, U_2, ..., U_n$ are exchangeable random variables, where U_j is defined as $U_j = sign(Y_j - M)$, $j = 1,2,...,n$. The proof to this follows from Theorem 2 by setting $g(\underline{X}) = g(X_1, ..., X_m) = M$ and $W_j = h(Y_j, g(X_1, ..., X_m)) = sign(Y_j - M)$.

The next step is to prove that when $F = G$, the joint distribution of $U_1, U_2, ..., U_n$ is distribution-free. To prove this we need to keep two things in mind. The first being that $P(U_j = 1 \text{ or } 0) = \frac{1}{2}$ since $(Y_j - M)$ is symmetric about zero when $F = G$. The second fact to recall is that $U_1, U_2, ..., U_n$ are exchangeable when $F = G$. The proof follows straightforwardly by combining these two facts.

In addition, when $F = G$, $U_j = sign(Y_j - M)$ and $V_j = |Y_j - M|$ for $j = 1,2,...,n$, are independent random variables. The proof follows from Lemma 1, since the distribution of $(Y_j - M)$ is symmetric about zero when $F = G$.

Next, we define $\underline{R}^* = (R_1^*, R_2^*, ..., R_n^*)$ where $R_j^* = 1 + \sum_{k=1}^{n} I(V_k < V_j)$

$= 1 + \sum_{k=1}^{n} I(|Y_k - M| < |Y_j - M|)$ for $j = 1,2,...,n$ (note that R_j^* is directly comparable to R_{ij}^* in equation (5.1)). Therefore, $\underline{R}^*$ is the vector of ranks of $V_1, V_2, ..., V_n$, i.e. $\underline{R}^*$ is the vector of ranks of $|Y_1 - M|, |Y_2 - M|, ..., |Y_n - M|$. We can prove, using Corollaries 1 and 2, that when $F = G$, any statistic that depends on the observations only through $U_1, U_2, ..., U_n$, i.e. $sign(Y_1 - M), sign(Y_2 - M), ..., sign(Y_n - M)$, and $\underline{R}^*$ is distribution-free over the class of continuous symmetric distributions. Consequently, the statistic $SRL_i = \sum_{j=1}^{n} sign(y_{ij} - M) R_{ij}^*$ is distribution-free. Since SRL_i is now known to be distribution-free, so is the signed-rank-like chart.

5.1.4. Simulation study

Bakir (2006) performed a simulation study where the robustness of the standard Shewhart $\overline{X}$ chart and that of the proposed Shewhart signed-rank-like chart are compared using the contaminated normal distribution. The contaminated normal distribution has been considered by various authors in an SPC context (see, for example, Wu, Zhao and Wang (2002) and Sheu and Yang (2006)). The cdf of the contaminated normal distribution is given by

$$\Phi_p(\theta,\sigma^2) = (1-p)\Phi(\theta,1) + p\Phi(\theta,\sigma^2) \tag{5.3}$$

where $(0 \leq)\ p\ (\leq 1)$ denotes the percentage of contamination, σ^2 (>0) denotes the severity of contamination and Φ denotes the cdf of the normal distribution, respectively. It should be noted that if $p = 0$ and $\theta = 0$ equation (5.3) reduces to the standard normal distribution. Bakir (2006) proved, through simulation, that if a process is contaminated by outliers it is ill-advised to use the standard Shewhart $\overline{X}$ chart, especially if the percentage of contamination (p) and/or the severity of contamination (σ^2) is high, i.e. $p > 0.01$ and/or $\sigma^2 > 4$. Bakir concludes that the Shewhart $\overline{X}$ chart is not robust against outliers, whereas the proposed Shewhart signed-rank-like chart is robust against outliers for all possible combinations of (p,σ^2). This is what we expected to find: the Shewhart signed-rank-like chart wouldn't be affected by outliers, since the median from the reference sample, the signs from the test sample and the ranks from the test sample aren't affected by outliers (recall that $SRL_i = \sum_{j=1}^{n} sign(y_{ij} - M)R_{ij}^{*}$).

Table 5.3 shows the simulated values of the ARL_0's of the two-sided Shewhart $\overline{X}$ chart for all possible combinations of (p,σ^2) with p =0.01, 0.05, 0.10, 0.15, 0.20, 1 and σ^2 = 4, 9, 16. These values are graphically illustrated in Figure 5.2. These simulated values are for a stable process with the presense of sporadic outliers. The case where the process is operational with no outliers, i.e. $p = 0$, is also given for reference. In these simulation studies 500 reference samples, each of size $m = 39$, were generated from the standard normal distribution. In addition, 500 test samples, each of size $n = 10$, were generated from the contaminated normal distribution.

Table 5.3. Simulated values of the ARL_0's for the two-sided Shewhart $\overline{X}$ chart*.

$LCL / UCL = \pm 2.8$				Severity of contamination		
				$\sigma^2 = 4$	$\sigma^2 = 9$	$\sigma^2 = 16$
				Level of severity		
				Low	Moderately high	High
Percentage of contamination	$p = 0$ (0%)	Level of p	None	163	163	163
	$p = 0.01$ (1%)		Low	159	115	70
	$p = 0.05$ (5%)		Moderately high	90	39	21
	$p = 0.10$ (10%)			61	22	12
	$p = 0.15$ (15%)		High	41	15	8
	$p = 0.20$ (20%)			33	11	6

Intuitively, we would expect the ARL to decrease (which would lead to an increase in the number of false alarms) as the percentage and/or severity of contamination increases. This is evident by looking at the lowest (p,σ^2) combination, i.e. $(p,\sigma^2) = (0, 4)$, opposed to the highest (p,σ^2) combination, i.e. $(p,\sigma^2) = (0.20, 16)$. The former shows that the ARL equals 163 when the process is operational with no outliers, whereas the latter shows that the ARL equals 6 when both the percentage and severity of contamination are high. These numbers indicate that there should be about 27 times as many false alarms when $(p,\sigma^2) = (0.20, 16)$ as opposed to $(p,\sigma^2) = (0,4)$.

Next, we look at what happens when both p and σ^2 are low. This is done by looking at the $(p,\sigma^2) = (0.01, 4)$ combination compared to the $(p,\sigma^2) = (0, 4)$ combination. The latter shows that the *ARL* equals 163 when the process is operational with no outliers, whereas the former shows that the *ARL* equals 159 when both the percentage and severity of contamination are low. These numbers indicate that there should be about the same number of false alarms when $(p,\sigma^2) = (0.01, 4)$ as opposed to $(p,\sigma^2) = (0,4)$.

* Table 5.3 appears in Bakir (2006), page 751, Table 1.

Next, we look at what happens when p is low, but σ^2 is moderately high. This is done by looking at the $(p,\sigma^2) = (0.01, 9)$ combination where the *ARL* has dropped to 115. This indicates that there should be about 1.42 times as many false alarms as the expected *ARL* of 163. Subsequently, we look at what happens when p is moderately high, but σ^2 is low. This is done by looking at the $(p,\sigma^2) = (0.05, 4)$ combination where the *ARL* has dropped to 90. This indicates that there should be about 1.81 as many false alarms as the expected *ARL* of 163. The rest of table can be interpreted similarly. The main conclusion that can be drawn from Table 5.3 is that it is ill-advised to use the Shewhart $\overline{X}$ chart when a process is contaminated by outliers, especially if the percentage of contamination (p) and/or the severity of contamination (σ^2) is high, i.e. $p > 0.01$ and/or $\sigma^2 > 4$.

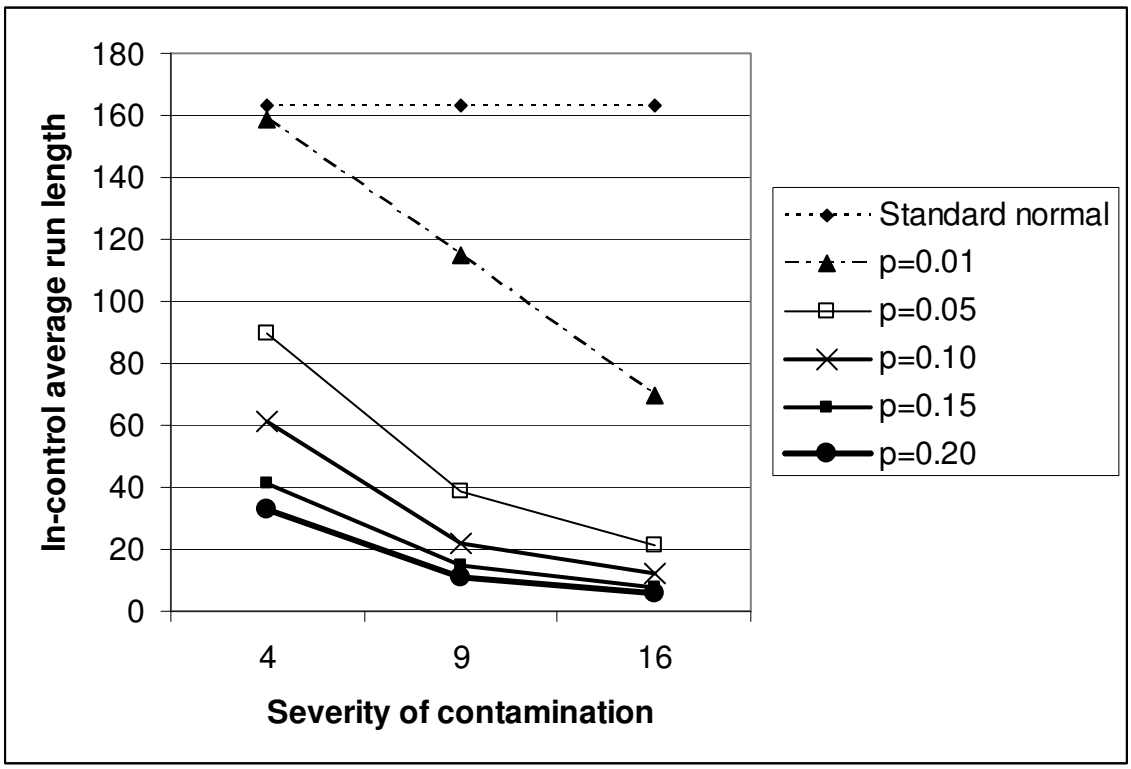

Figure 5.2. Simulated ARL_0 values for the two-sided Shewhart $\overline{X}$ chart for various values of p and σ^2.

5.1.5. Comparisons

The first comparison between the standard Shewhart $\overline{X}$ chart and the proposed Shewhart signed-rank-like chart.

Table 5.4. Simulated values of the *ARL* for the two-sided Shewhart $\overline{X}$ control chart ($\overline{X}_{CC}$) and the Shewhart signed-rank-like control chart (SRL_{CC})*.

	Severity of contamination											
	$\sigma^2 = 4$				**$\sigma^2 = 9$**				**$\sigma^2 = 16$**			
	Level of severity											
	Low				**Moderately high**				**High**			
	$p = 0.01$ (1%)		**$p = 0.10$ (10%)**		**$p = 0.01$ (1%)**		**$p = 0.10$ (10%)**		**$p = 0.01$ (1%)**		**$p = 0.10$ (10%)**	
Control chart	$\overline{X}_{CC}$	SRL_{CC}	$\overline{X}_{CC}$	SRL_{CC}	$\overline{X}_{CC}$	SRL_{CC}	$\overline{X}_{CC}$	SRL_{CC}	$\overline{X}_{CC}$	SRL_{CC}	$\overline{X}_{CC}$	SRL_{CC}
LCL / UCL	**± 2.85**	**± 53**	**± 3.25**	**± 53**	**± 2.96**	**± 53**	**± 3.98**	**± 53**	**± 3.24**	**± 53**	**± 4.92**	**± 53**
$\theta = 0.0$	170.0	166.0	165.0	166.0	167.4	166.0	166.7	166.0	164.6	166.0	167.0	166.0
$\theta = 0.2$	122.8	121.8	131.8	128.3	115.4	120.4	124.3	122.2	137.3	130.2	145.6	128.6
$\theta = 0.4$	43.9	50.9	59.8	54.4	56.4	60.7	65.4	66.2	60.6	61.7	88.3	65.0
$\theta = 0.6$	10.6	24.2	13.5	27.2	14.4	23.7	27.5	26.5	17.7	22.1	42.8	31.5
$\theta = 0.8$	3.5	11.1	5.4	10.6	4.2	9.3	11.9	10.3	6.7	8.4	22.6	11.3
$\theta = 1.0$	1.9	4.4	2.8	5.5	2.1	4.7	5.3	5.6	3.5	4.5	13.4	5.9

* Table 5.4 appears in Bakir (2006), page 754, Table 3.

Bakir (2006) compared the proposed Shewhart signed-rank-like chart to the Shewhart $\overline{X}$ chart using the contaminated normal distribution (the observations are normally distributed with occasional outliers). Both charts are designed to have approximately the same in-control average run length to ensure fair comparison between the charts. The out-of-control average run lengths were computed, using these chart constants, for various values of the median θ, the percentage of contamination (p) and the severity of contamination (σ^2). We typically want the ARL_δ to be small, i.e. the chart with the smallest ARL_δ will be the preferred chart.

From Table 5.4 we see that the median ranges from 0 (the in-control value) to 1 in increments of 0.2; the severity of contamination ranges from low to high, that is, $\sigma^2 = 4$ (low), $\sigma^2 = 9$ (moderately high) and $\sigma^2 = 16$ (high); and the percentage of contamination is taken to be 1% (low) and 10% (moderately high), respectively.

We start by investigating the lowest percentage and severity of contamination levels for the smallest process shift of 0.2. The ARL_δ of the Shewhart $\overline{X}$ chart (=122.8) is almost equivalent to the ARL_δ of the Shewhart signed-rank-like chart (=121.8). Therefore, for a low percentage and severity of contamination and a small process shift, both charts are performing equally well. More generally, for low to moderately high levels of p (= 0.01 or 0.1) and σ^2 (= 4 or 9) and small process shifts ($\theta = 0.2$ or 0.4), the ARL_δ values of the Shewhart signed-rank-like chart are almost equivalent to the ARL_δ values of the Shewhart $\overline{X}$ chart.

In contrast, we investigate the highest percentage and severity of contamination for the largest process shift of 1. The ARL_δ of the Shewhart $\overline{X}$ chart (=13.4) is higher than the ARL_δ of the Shewhart signed-rank-like chart (=5.9). Consequently, we see that the Shewhart signed-rank-like chart performs better than the Shewhart $\overline{X}$ chart for a high percentage and severity of contamination and a large process shift. More generally, we find that for $p = 0.1$ and $\sigma^2 = 16$ the ARL_δ values of the Shewhart $\overline{X}$ chart are *all* higher than the ARL_δ values of the Shewhart signed-rank-like chart for *all* process shifts ($\theta = 0.2$, 0.4, 0.6, 0.8 and 1). As a result we conclude that the Shewhart signed-rank-like chart performs better than the Shewhart $\overline{X}$ chart for high levels of p and σ^2 over all process shifts.

It should be noted that there are various cases where the Shewhart $\overline{X}$ chart performs better than the Shewhart signed-rank-like chart. An illustration of the latter is, for example, for low to moderately high levels of p (= 0.01 or 0.1), $\sigma^2 = 4$ and large process shifts ($\theta = 0.6$, 0.8 and 1) the ARL_δ values of the Shewhart $\overline{X}$ chart are lower than those of the Shewhart signed-rank-like chart. Hence, in some cases the Shewhart $\overline{X}$ chart outperforms the Shewhart signed-rank-like chart and vice versa. Table 5.5 indicates which chart, between the Shewhart $\overline{X}$ chart and the proposed signed-rank-like chart, is the preferred chart for various values of p, σ^2 and θ. The term 'comparable' in Table 5.5 implies that the proposed signed-rank-like chart is as efficient as the Shewhart $\overline{X}$ chart.

Table 5.5. Summary of the first comparison between the Shewhart $\overline{X}$ chart and the proposed signed-rank-like chart*.

	$\sigma^2 = 4$	$\sigma^2 = 9$	$\sigma^2 = 16$
$p = 0.01$	Small shifts: Comparable Large shifts: $\overline{X}$	Small shifts: Comparable Large shifts: $\overline{X}$	Small shifts: Comparable Large shifts: $\overline{X}$
$p = 0.10$	Small shifts: Comparable Large shifts: $\overline{X}$	All shifts: Comparable	All shifts: Signed-rank-like chart

The second comparison between the standard Shewhart $\overline{X}$ chart and the proposed Shewhart signed-rank-like chart.

The out-of-control *ARL* is examined for three distributions, namely, the Normal, Laplace and Cauchy distributions, respectively. Recall that we want the ARL_δ to be small in all cases.

* Small shifts refer to $\theta = 0.2$ or 0.4, whereas large shifts refer to $\theta = 0.6, 0.8$ or 1.

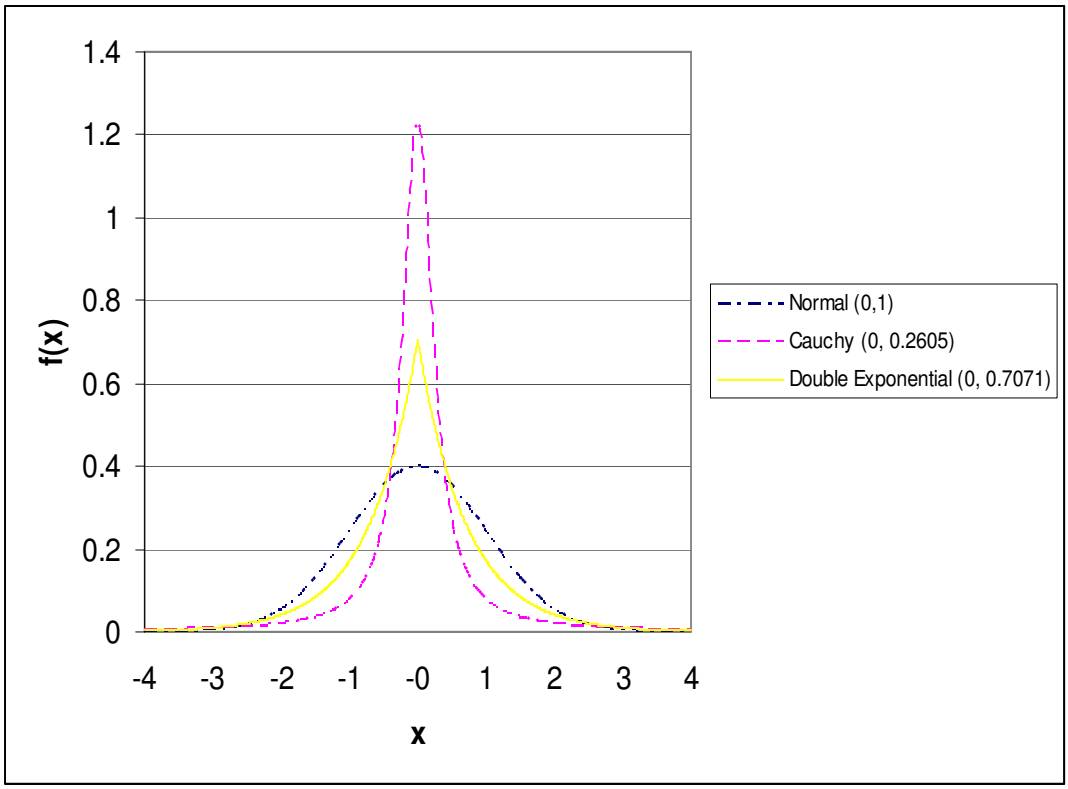

Figure 5.3. The shapes of the three distributions under consideration.

(i) The Normal distribution

For the Normal distribution we would expect the out-of-control performance of the Shewhart $\overline{X}$ chart to be better than that of the Shewhart signed-rank-like chart. The chart constants for both the Shewhart signed-rank-like and Shewhart $\overline{X}$ charts are chosen such that the in-control average run length is approximately equal ($ARL_0 \approx 164$) for both charts: $LCL/UCL_{\overline{X}} = \pm 2.80$ and $LCL/UCL_{SRL} = \pm 53$. The out-of-control average run length values were computed, using these chart constants, for various values of the median θ. The median ranges from 0 (the in-control value) to 1 in increments of 0.2. The results are shown below in Figure 5.4.

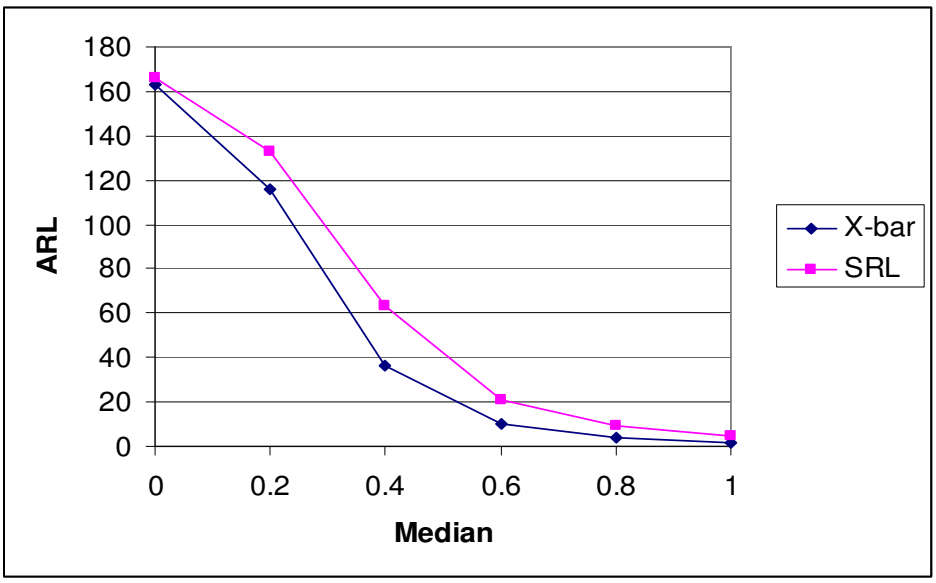

Figure 5.4. Comparison of the Shewhart signed-rank-like chart with the Shewhart $\overline{X}$ chart under Normal shift alternatives.

When comparing the Shewhart signed-rank-like chart with the Shewhart $\overline{X}$ chart under Normal shift alternatives we find that the Shewhart $\overline{X}$ chart is performing better than the Shewhart signed-rank-like chart, since the out-of-control average run length values for the Shewhart $\overline{X}$ chart are smaller than the out-of-control average run length values for the Shewhart signed-rank-like chart. However, it should be noted that the differences are small and it appears to fade away when the process is shifted from its in-control value of 0 to values greater than 0.8.

(ii) The Double Exponential distribution

The Double Exponential distribution, also called the Laplace distribution, is comparable to the Normal distribution (since they are both symmetric around 0), but it has heavier tails (see Figure 5.3). As a result, there are higher probabilities associated with extreme values when working with the Double Exponential distribution as opposed to using the Normal distribution. The scale parameter λ of the Double Exponential distribution is set equal to $1/\sqrt{2}$ so that the Double Exponential distribution has a standard deviation of 1. For the Double Exponential distribution we would expect the out-of-control performance of the Shewhart signed-rank-like chart to be better than that of the Shewhart $\overline{X}$ chart. The chart constants for both the Shewhart signed-rank-like and Shewhart $\overline{X}$ chart are chosen such that

the in-control average run length is approximately equal ($ARL_0 \approx 150$) for both charts: $LCL/UCL_{\bar{X}} = \pm 2.85$ and $LCL/UCL_{SRL} = \pm 53$. The results are shown below in Figure 5.5.

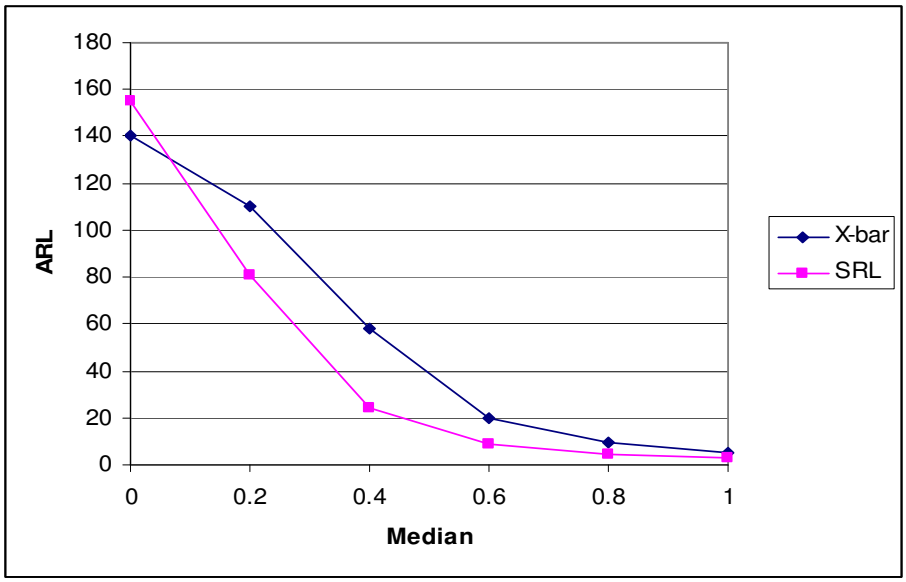

Figure 5.5. Comparison of the Shewhart signed-rank-like chart with the Shewhart $\bar{X}$ chart under Double Exponential shift alternatives.

When comparing the Shewhart signed-rank-like chart with the Shewhart $\bar{X}$ chart under Double Exponential shift alternatives we find that the Shewhart signed-rank-like chart is performing better than the Shewhart $\bar{X}$ chart, since the out-of-control average run length values for the Shewhart signed-rank-like chart are smaller than the out-of-control average run length values for the Shewhart $\bar{X}$ chart. However, it should be noted that the differences are small and it appears to fade away when the process is shifted from its in-control value of 0 to values greater than 0.8.

(iii) The Cauchy distribution

The scale parameter λ of the Cauchy distribution is set equal to 0.2605 so that the Cauchy distribution has a probability of 0.95 to the left of 1.645 (which is also the case for the standard normal distribution). For the Cauchy distribution we would expect the out-of-control performance of the Shewhart signed-rank-like chart to be better than that of the Shewhart $\bar{X}$ chart. The chart constants for both the Shewhart signed-rank-like and Shewhart $\bar{X}$ chart are chosen such that the in-control average run length is approximately equal ($ARL_0 \approx 164$) for

both charts: $LCL/UCL_{\bar{X}} = \pm 22$ and $LCL/UCL_{SRL} = \pm 53$. The results are shown below in Figure 5.6.

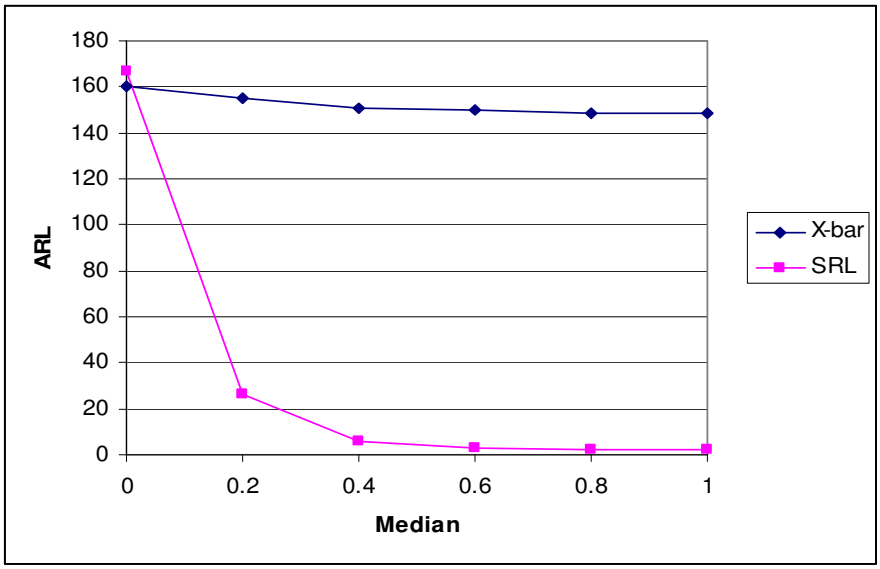

Figure 5.6. Comparison of the Shewhart signed-rank-like chart with the Shewhart $\bar{X}$ chart under Cauchy shift alternatives.

When comparing the Shewhart signed-rank-like chart with the Shewhart $\bar{X}$ chart under Cauchy shift alternatives we find that the Shewhart signed-rank-like chart is performing better than the Shewhart $\bar{X}$ chart, since the out-of-control average run length values for the Shewhart signed-rank-like chart are smaller than the out-of-control average run length values for the Shewhart $\bar{X}$ chart. It should be noted that these differences are large for all values of the median θ.

In conclusion we found that the Shewhart signed-rank-like chart performs better than the Shewhart $\bar{X}$ chart under heavy tailed distributions. In addition, recall that the Shewhart $\bar{X}$ chart is not robust against outliers, whereas the proposed Shewhart signed-rank-like chart is, for the most part, robust against outliers. These are two key motivations to why the user should rather use the Shewhart signed-rank-like chart as opposed to using the Shewhart $\bar{X}$ chart.

Table 5.6. Summary of the second comparison between the Shewhart $\overline{X}$ chart and the proposed signed-rank-like chart.

Distribution	**Preferred control chart**
Normal	Shewhart $\overline{X}$ chart
Double Exponential	Shewhart signed-rank-like chart
Cauchy	Shewhart signed-rank-like chart

5.1.6. The tabular CUSUM control chart

Bakir (2006) proposed a tabular CUSUM signed-rank-like chart. Generally, the standardized upper one-sided CUSUM is given by

$$S_i^+ = \max[0, S_{i-1}^+ + y_i - k] \quad \text{for } i = 1,2,3,... \tag{5.4}$$

while the resulting standardized lower one-sided CUSUM is given by

$$S_i^- = \min[0, S_{i-1}^- + y_i + k] \quad \text{for } i = 1,2,3,... \tag{5.5}$$

or

$$S_i^{-*} = \max[0, S_{i-1}^{-*} - y_i - k] \quad \text{for } i = 1,2,3,... \tag{5.6}$$

The two-sided standardized CUSUM is constructed by running the upper and lower one-sided standardized CUSUM charts simultaneously and signals at the first i such that $S_i^+ \geq h$ or $S_i^- \leq -h$.

The chart proposed by Bakir (2006) instead uses the cumulative sum of the statistic SRL_i (defined in (5.2)) with a stopping rule. A CUSUM signed-rank-like chart can be obtained by replacing y_i in expressions (5.4), (5.5) and (5.6) with SRL_i. In other words, for the upper one-sided CUSUM signed-rank-like chart we use

$$S_i^+ = \max[0, S_{i-1}^+ + SRL_i - k] \quad \text{for } i = 1,2,3,... \tag{5.7}$$

to detect positive deviations from zero. A signalling event occurs for the first i such that $S_i^+ \geq h$.

For a lower one-sided CUSUM signed-rank-like chart we use

$$S_i^- = \min[0, S_{i-1}^- + SRL_i + k] \quad \text{for } i = 1,2,3,... \tag{5.8}$$

or

$$S_i^{-*} = \max[0, S_{i-1}^{-*} - SRL_i - k] \quad \text{for } i = 1,2,3,... \tag{5.9}$$

to detect negative deviations from zero. A signalling event occurs for the first i such that $S_i^- \leq -h$ (if expression (5.8) is used) or $S_i^{-*} \geq h$ (if expression (5.9) is used).

The corresponding two-sided CUSUM chart signals for the first i at which either one of the two inequalities is satisfied, that is, either $S_i^+ \geq h$ or $S_i^- \leq -h$. Starting values are typically chosen to equal zero, that is, $S_0^+ = S_0^- = 0$.

A CUSUM signed-rank-like chart can also be constructed by replacing y_i in expressions (5.4), (5.5) and (5.6) with the standardized signed-rank-like statistic.

Although Bakir (2006) provided the general idea of how to construct a CUSUM signed-rank-like control chart, he failed to do any simulation studies or to give any tables that can be used for the implementation of the chart. More research is necessary on CUSUM signed-rank-like control charts, for example, one could look at the implementation of the CUSUM signed-rank-like chart and study its performance.

5.1.7. The EWMA control chart

Bakir (2006) proposed an EWMA signed-rank-like chart. Generally, an EWMA control chart scheme accumulates statistics $X_1, X_2, X_3, \ldots$ with the plotting statistics defined as

$$Z_i = \lambda X_i + (1-\lambda) Z_{i-1} \tag{5.10}$$

where $0 < \lambda \leq 1$ is a constant called the weighting constant. The starting value Z_0 is often taken to be zero.

A nonparametric EWMA-type of control chart based on the signed-rank-like statistic can be obtained by replacing X_i in expression (5.10) with SRL_i. Therefore, the EWMA signed-rank-like chart accumulates the statistics $SRL_1, SRL_2, SRL_3, \ldots$ with the plotting statistics defined as

$$Z_i = \lambda SRL_i + (1-\lambda) Z_{i-1} \tag{5.11}$$

where $0 < \lambda \leq 1$ and the starting value Z_0 could be taken to equal zero, i.e. $Z_0 = 0$.

An EWMA signed-rank-like chart can also be constructed by replacing X_i in expression (5.10) with the standardized signed-rank-like statistic.

Although Bakir (2006) provided the general idea of how to construct an EWMA signed-rank-like control chart, he failed to do any simulation studies or to give any tables that can be used for the implementation of the chart. More research is necessary on EWMA signed-rank-like control charts, for example, one could look at the implementation of the EWMA signed-rank-like chart and study its performance.

5.1.8. Summary

In this chapter we examined the Shewhart-type signed-rank-like chart proposed by Bakir (2006). We illustrated these procedures using the piston ring data from Montgomery (2001) to help the reader to understand the subject more thoroughly. The proposed chart is recommended when the process distribution is known to be heavy-tailed or to be contaminated by occasional outliers. We also briefly looked at CUSUM- and EWMA-type signed-rank-like charts. Although Bakir (2006) provided general ideas on how to construct CUSUM- and EWMA-type signed-rank-like control charts, he failed to do any simulation studies or to give any tables that can be used for the implementation of these charts. More research is necessary on CUSUM- and EWMA-type signed-rank-like control charts, for example, one could look at the implementation of these charts and study their performance.

Chapter 6: Mann-Whitney-Wilcoxon control charts

6.1. The Shewhart-type control chart

6.1.1. Introduction

While the precedence chart is a step in the right direction, the precedence test is not the most popular or the most powerful of nonparametric two-sample tests. That honour goes to the Mann-Whitney test (see for example, Gibbons and Chakraborti, 2003). The Mann-Whitney (hereafter MW) test (equivalent to the popular Wilcoxon rank-sum test) is a well-known nonparametric competitor of the two-independent-sample *t*-test. The test is known to be more powerful than the precedence test for light tailed distributions and hence MW charts are expected to be more efficient for such cases. Of course, the MW chart is also distribution-free and therefore has the same in-control robustness advantage as the precedence chart, namely that its in-control distribution is completely known. Park and Reynolds (1987) considered Shewhart-type control charts for monitoring the location parameter of a continuous process in case U. One of the special cases of their charts is the MW chart based on the Mann-Whitney-Wilcoxon (hereafter MWW) statistic. The control limits of these charts are established using Phase I reference data. However, they only considered properties of this chart when the reference sample size approaches infinity. Chakraborti and Van de Wiel (2003) considered the Shewhart-type MW chart for finite reference sample size, studied its properties, and provided tables for its implementation. These authors show that in some cases the MW chart is more efficient than the precedence chart.

Assume that a reference sample of size m, $X_1, X_2, ..., X_m$, is available from an in-control process with an unknown continuous cdf $F(x)$. Let $Y_1^h, Y_2^h, ..., Y_{n_h}^h$, $h = 1,2,...,$ denote the h^{th} test sample of size n_h. Let $G^h(y)$ denote the cdf of the distribution of the h^{th} Phase II sample. $G^h(y) = G(y) \quad \forall h$, since the Phase II samples are all assumed to be identically distributed. Accordingly, the superscript h can be suppressed from this point forward. For convenience, assume that the Phase II samples are all of the same size, n. The Mann-Whitney test is based on the total number of (X, Y) pairs where the Y-observation (Phase II sample) is strictly greater than the X-observation (Phase I sample).

6.1.2. Plotting statistic

The Mann-Whitney statistic is defined to be

$$M_{XY} = \text{the number of pairs } (X_i, Y_j) \text{ with } Y_j > X_i \tag{6.1}$$

for $i = 1,2,...,m$ and $j = 1,2,...,n$. Expression (6.1) can be written as

$$M_{XY} = \sum_{i=1}^{m}\sum_{j=1}^{n} I(Y_j > X_i) \tag{6.2}$$

where $I(Y_j > X_i)$ is the indicator function, i.e.

$$I(Y_j > X_i) = \begin{cases} 1 & \text{if} \quad Y_j > X_i \\ 0 & \text{if} \quad Y_j \le X_i \end{cases}.$$

There are a total of mn (X_i, Y_j) pairs for each Phase II sample. Therefore, if all the Y-observations are greater than the X-observations, M_{XY} would be equal to mn. On the other hand, if all the Y-observations are smaller than the X-observations, M_{XY} would be equal to 0. Therefore, we have that $0 \le M_{XY} \le mn$. For large values of M_{XY}, that is, if a large number of the Y-observations are greater than the X-observations, this would be indicative of a positive shift from the X to the Y distribution. On the other hand, for small values of M_{XY}, that is, if a large number of the Y-observations are smaller than the X-observations, this would be indicative of a negative shift from the X to the Y distribution.

The proposed MW chart plots the M_{XY} statistics, that is, $M_{XY}^1, M_{XY}^2, \ldots$, versus the test sample number. M_{XY} is referred to as the plotting statistic. The chart signals if the plotting statistic falls on or above the upper control limit (UCL) or if the plotting statistic falls on or below the lower control limit (LCL). Since the in-control distribution of the plotting statistic, M_{XY}, is symmetric about the mean $\frac{mn}{2}$ (see Gibbons and Chakraborti (2003)), the control limits are taken to be symmetric. Because of symmetry, we have that $P(M_{XY} = a) = P(M_{XY} = mn - a)$ for the constant a with $0 \le a \le mn$, so it is reasonable to take $L_{mn} = mn - U_{mn}$ where U_{mn} and L_{mn} denote the upper and lower control limits, respectively. If the plotting statistic M_{XY} falls between the control limits, that is, $L_{mn} < M_{XY} < U_{mn}$, the process is declared to be in-control, whereas if the plotting statistic

M_{XY} falls on or outside one of the control limits, that is, if $M_{XY} \leq L_{mn}$ or $M_{XY} \geq U_{mn}$, the process is declared to be out-of-control.

6.1.3. Properties of the run-length distribution

Result 6.1: Probability of a signal - conditional

Let $p_G(\underline{x})$ denote the probability of a signal with any test (Phase II) sample, given the reference sample $(X_1, X_2, ..., X_m) = (x_1, x_2, ..., x_m)$ (in short, $\underline{X} = \underline{x}$).

$$p_G(\underline{x}) = 2P_G\left(M_{\underline{x}Y} \geq U_{mn}\right)$$

$$\begin{aligned} p_G(\underline{x}) &= P_G(\text{Signal} \mid \underline{X} = \underline{x}) \\ &= P_G\left(M_{\underline{x}Y} \leq L_{mn}\right) + P_G\left(M_{\underline{x}Y} \geq U_{mn}\right) \\ &= P_G\left(M_{\underline{x}Y} \leq mn - U_{mn}\right) + P_G\left(M_{\underline{x}Y} \geq U_{mn}\right) \\ &= 2P_G\left(M_{\underline{x}Y} \geq U_{mn}\right). \end{aligned}$$

The last equality follows on account of symmetry (see Section 6.1.2). From Result 6.1 it can be seen that the calculation of $p_G(\underline{x})$ essentially requires the calculation of the upper-tailed probability $P_G\left(M_{\underline{x}Y} \geq U_{mn}\right)$. More detail on this point appears in Section 6.1.4.

Result 6.2: Probability of no signal - conditional

$$P_G(\text{No Signal} \mid \underline{X} = \underline{x}) = 1 - p_G(\underline{x}) = 1 - 2P_G\left(M_{\underline{x}Y} \geq U_{mn}\right)$$

Result 6.3: Run-length distribution - conditional

$$P(N = k \mid \underline{X} = \underline{x}) = (1 - p_G(\underline{x}))^{k-1}(p_G(\underline{x})) \quad \text{for} \quad k = 1,2,3,...$$

The conditional run length, denoted by $N \mid \underline{X} = \underline{x}$, will have a geometric distribution with parameter $p_G(\underline{x})$, because all the Phase II samples are independent if we condition on the reference sample. A detailed motivation for using the method of conditioning is given by

Chakraborti (2000), but, in brief, the signalling events are dependent and by means of conditioning on the reference sample we don't have to be concerned about the dependence. Consequently we have that

$$N \mid \underline{X} = \underline{x} \sim GEO\left(p_G(\underline{x})\right)$$

$$P\left(N = k \mid \underline{X} = \underline{x}\right) = \left(1 - p_G(\underline{x})\right)^{k-1}\left(p_G(\underline{x})\right) \quad \text{for} \quad k = 1,2,3,...$$

Consequently, the cumulative distribution function (cdf) is found from

$$P\left(N \le k \mid \underline{X} = \underline{x}\right) = \sum_{i=1}^{k}\left(1 - p_G(\underline{x})\right)^{i-1}\left(p_G(\underline{x})\right) \quad \text{for} \quad k = 1,2,3,...$$

Result 6.4: Average run-length – conditional

$$CARL = E_G\left(N \mid \underline{X} = \underline{x}\right) = \frac{1}{p_G(\underline{x})}$$

Since the conditional run length, denoted by $N \mid \underline{X} = \underline{x}$, has a geometric distribution with parameter $p_G(\underline{x})$, the conditional average run length is given by

$$CARL = E_G\left(N \mid \underline{X} = \underline{x}\right) = \frac{1}{p_G(\underline{x})}.$$

Result 6.5: Average run-length – unconditional

$$UARL = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \upsilon\left(G(x_1), G(x_2), ..., G(x_m)\right) dF(x_1) \ldots dF(x_m)$$

where

υ is some function of G and $x_1, x_2, ..., x_m$.

$$\begin{aligned} & UARL \\ &= E(N) \\ &= E_F\left(E_G(N \mid \underline{X} = \underline{x})\right) \\ &= E_F\left(\frac{1}{p_G(\underline{x})}\right) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{p_G(\underline{x})} dF(x_1) \ldots dF(x_m). \end{aligned} \tag{6.3}$$

The second equality in (6.3) follows from extending the notion of expectation to the conditional framework. The third equality in (6.3) follows from Result 6.4. The fourth equality in (6.3) follows from the definition of expected values (see, for example, Bain and Engelhardt (1992)).

From (6.3) it can be seen that the unconditional ARL is an m-dimensional integral, since the reference sample is of size m. Equation (6.3) can be expressed differently by writing $\frac{1}{p_G(\underline{x})}$ as $\upsilon\left(P(Y_j < x_1), P(Y_j < x_2),..., P(Y_j < x_m)\right) = \upsilon\left(G(x_1), G(x_2),..., G(x_m)\right)$, where υ is some function of G and $x_1, x_2,..., x_m$. By substituting $\frac{1}{p_G(\underline{x})}$ in (6.3) with $\upsilon\left(G(x_1), G(x_2),..., G(x_m)\right)$ we obtain

$$UARL = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \upsilon\left(G(x_1), G(x_2),..., G(x_m)\right) dF(x_1) \ldots dF(x_m). \tag{6.4}$$

Recall that a process is said to be in-control when $G = F$. Therefore, the in-control (unconditional) ARL is obtained by substituting $G = F$ into the equation for the unconditional ARL given in (6.3) and we obtain the m-dimensional integral

$$UARL_0 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{p_F(\underline{x})} dF(x_1) \ldots dF(x_m), \tag{6.5}$$

where the subscript 0 refers to the in-control state.

In the out-of-control case the unconditional ARL is given by the m-dimensional integral

$$UARL_\delta = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{p_G(\underline{x})} dF(x_1) \ldots dF(x_m), \tag{6.6}$$

where δ signifies a shift between F and G.

Recall that $\frac{1}{p_G(\underline{x})}$ was re-written as $\upsilon\left(G(x_1), G(x_2), ..., G(x_m)\right)$ where υ is some function of G and $x_1, x_2, ..., x_m$. Similarly, $\frac{1}{p_F(\underline{x})}$ can be re-written as $\upsilon\left(F(x_1), F(x_2), ..., F(x_m)\right)$ and we obtain

$$
\begin{aligned}
&UARL_0 \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \upsilon\left(F(x_1), F(x_2), ..., F(x_m)\right) dF(x_1) \ldots dF(x_m) \\
&= \int_{0}^{1} \cdots \int_{0}^{1} \upsilon(u_1, u_2, ..., u_m) du_1 \ldots du_m \\
&= \int_{0}^{1} \cdots \int_{0}^{1} \frac{1}{p_U(\underline{u})} du_1 \ldots du_m, \qquad (6.7)
\end{aligned}
$$

by the probability integral transformation (see, for example, Gibbons and Chakraborti (2003)). The subscript U refers to the uniform(0,1) distribution and $p_U(\underline{u})$ is the conditional probability of a signal at any test sample, given the reference sample, when the process is in-control.

Recall that for the in-control case, the distributions of both the reference and test samples can be assumed to be uniformly(0,1) distributed*, which shows that the unconditional ARL, for the in-control situation, of the MW chart does not depend on the underlying process distributions F and G. The same argument can be used to show that the in-control run length distribution does not depend on the underlying process distributions F and G, thus establishing that the proposed MW chart is distribution-free.

We have to calculate the unconditional ARL, for the in-control situation, using (6.7) to implement the chart. Following this, we have to calculate the unconditional ARL, for the out-of-control situation, using (6.3) to evaluate chart performance. We run into two problems in doing so, that is, (i) we don't have exact formulas for the signal probabilities $p_G(\underline{x})$ and $p_U(\underline{u})$; and (ii) it could be difficult and time-consuming estimating (6.3) and (6.7), since both

* For the in-control case, the distributions of both the reference and test samples can be assumed to be uniformly(0,1) distributed. This is due to the well-known probabitliy integral transformation (see, for example, Gibbons and Chakraborti (2003)).

unconditional average run length formulas (for the in-control and out-of-control situations, respectively) are m-dimensional integrals.

Chakraborti and Van de Wiel (2003) proposed a possible solution to both of these problems. Their proposed solution proceeds in two steps. Firstly, fast computations (or approximations) of the signal probabilities are done. It should be noted that although the computation of $p_G(\underline{x})$ will be discussed in detail (in the following section), the computation of $p_U(\underline{u})$ is omitted, since it follows similarly to the computation of $p_G(\underline{x})$. Secondly, Monte Carlo simulation is applied to approximate the unconditional ARL's (for the in-control and out-of-control situations, respectively). The Monte Carlo estimates are given by

$$A\hat{R}L \approx \frac{1}{K}\sum_{i=1}^{K}\frac{1}{p_G(\underline{x}_i)} \tag{6.8}$$

and

$$A\hat{R}L_0 \approx \frac{1}{K}\sum_{i=1}^{K}\frac{1}{p_U(\underline{u}_i)} \tag{6.9}$$

where K denotes the number of Monte Carlo samples, $\underline{x}_i = (x_{i1}, x_{i2}, ..., x_{im})$ and $\underline{u}_i = (u_{i1}, u_{i2}, ..., u_{im})$ denote the i^{th} Monte Carlo sample, $i = 1,2,...,K$, of which each element is taken from some specified F for the $A\hat{R}L$ (for the out-of-control situations) and from the uniform(0,1) distribution for the $A\hat{R}L_0$ (for the in-control situation).

One concern is the size of K, that is, how may Monte Carlo samples should be used? Although larger sizes of K can result in more accurate approximations and smaller Monte Carlo errors, using larger Monte Carlo samples may be more time-consuming. This concern will be addressed in Section 6.1.6.

6.1.4. The computation of the signal probability

The Mann-Whitney statistic, given in (6.2), can be written in a simpler (more straightforward) form given by $M_{\underline{x}\underline{Y}} = \sum_{j=1}^{n} C_j$, where C_j denotes the number of x-observations that precede Y_j, $j = 1,2,...,n$. Also recall that since $p_G(\underline{x}) = 2P_G\left(M_{\underline{x}\underline{Y}} \geq U_{mn}\right)$,

the calculation of $p_G(\underline{x})$ essentially requires the calculation of the upper-tailed probability $P_G\left(M_{\underline{x}Y} \geq U_{mn}\right)$. The computation of the latter proceeds in two steps, namely: (i) listing of all n-tuples $(C_1, C_2, ..., C_n)$ for which the sum is greater than or equal to U_{mn}; and (ii) the summation of the probabilities for these tuples.

The Central Limit Theorem states that if S_n is the sum of n variables, then the distribution of S_n approaches the normal distribution as n approaches infinity, i.e. $S_n \rightarrow$ Normal distribution as $n \rightarrow \infty$. Using this result, we can find a normal approximation to the upper-tailed probability $P_G\left(M_{\underline{x}Y} \geq U_{mn}\right)$, since $M_{\underline{x}Y} = \sum_{j=1}^{n} C_j$ approaches the normal distribution as $n \rightarrow \infty$. Although using a normal approximation to the upper-tailed probability is a possible solution, it is not ideal. The reason being that although normal approximations work well when n is large (and improve as sample size increases), normal approximations do not work well when n is considered small. In our applications we typically use sample sizes that may be considered small and as a result using normal approximations would be somewhat unattractive. Clearly, a better approach is needed.

6.1.5. Saddlepoint approximations

Saddlepoint approximations (or saddlepoint expansions) provide good approximations (with a small relative error) to very small tail probabilities. Consequently, saddlepoint approximations can be applied to the problem of finding $p_G(\underline{x})$, which is usually set to be rather small (typically 0.0027). Jensen (1995) provides ample justifications for the application of saddlepoint expansions when approximating small probabilities. In Chapter 2 of Jensen (1995) the classical saddlepoint approximations for tail probabilities for sums of independent random variables are given. For our problem (the calculation of the upper-tailed probability $P_G(M_{\underline{x}Y} \geq U_{mn})$), we make use of the Lugannani-Rice formula (hereafter LR-formula) which is a saddlepoint expansion formula.

Prior to defining the LR-formula, a few concepts will be explained. To begin with, let a_l denote the probability that l x-observations (given $\underline{X} = \underline{x}$) precede Y_j for $j = 1,2,...,n$

and $l = 0,1,...,m$, respectively. Therefore, $a_l = P(C_j = l \mid \underline{X} = \underline{x}) = P(x_{(l)} < Y_j \le x_{(l+1)})$ where $x_{(1)} \le x_{(2)} \le \ldots \le x_{(m)}$ are the order statistics.

Since the pgf provides a very useful tool for studying the sum of independent random variables we turn to the conditional probability generating function (pgf) of C_j, and subsequently to the conditional pgf of $M_{\underline{x}Y}$. In view of the fact that C_j, $j = 1,2,...,n$, is a random variable whose possible values are restricted to the nonnegative integers $\{0,1,...,m\}$, the conditional pgf of C_j is given by

$$\Pi_1(z) = \sum_{l=0}^{m} P(C_j = l \mid \underline{X} = \underline{x}) z^l = \sum_{l=0}^{m} a_l z^l \ . \tag{6.10}$$

It's a well-known fact that if, for example, X and Y are independent random variables with probability generating functions $\Pi_X(z)$ and $\Pi_Y(z)$, respectively, we have that

$$\Pi_{X+Y}(z) = \Pi_X(z)\Pi_Y(z) \tag{6.11}$$

(see, for example, Bain and Engelhardt (1992)).

$M_{\underline{x}Y}$ is the sum of n independent identical variables (recall that $M_{\underline{x}Y} = \sum_{j=1}^{n} C_j$) and therefore, by using (6.10) and (6.11), the conditional pgf of $M_{\underline{x}Y}$ is given by

$$\Pi_2(z) = \sum_{j=0}^{mn} P(M_{\underline{x}Y} = j) z^j = \left(\sum_{j=0}^{m} a_j z^j \right)^n \ . \tag{6.12}$$

By implication C_j, for $j = 1,2,...,n$, are independent identically distributed, conditionally.

Next we examine the cumulant generating function (cgf) of C_j. The cgf is just the logarithm of the moment generating function (mgf). Mathematically, the mgf and the cgf are equivalent. The cgf generates the mean and variance, instead of the uncentered moments. We can think of $\kappa'(t)$ and $\kappa''(t)$ as the mean and variance, respectively, where $\kappa(t)$ denotes the cgf. Hence, the cgf of C_j can be obtained by taking the logarithm of the pgf in (6.12) at the point $z = e^t$. As a result, the cgf of C_j is given by

$$\kappa(t) = \log\left(\left(\sum_{l=0}^{m} a_l z^l\right)_{z=e^t}\right) = \log\left(\sum_{l=0}^{m} a_l e^{tl}\right). \tag{6.13}$$

The first and second order derivatives of the cgf is simply $\kappa'(t)$ and $\kappa''(t)$ so that $\kappa'(t) = m(t)$ and $\kappa''(t) = \sigma^2(t)$. Also, let $\mu = \frac{U_{mn}}{n}$ and $\overline{M}_{\underline{x}Y} = \frac{M_{\underline{x}Y}}{n}$. The saddlepoint, γ, is the solution to the equation $m(t) = \mu$. In other words, we solve $m(t) = \mu$ for t (see Theorem 3 in Appendix A for a detailed discussion on saddlepoint techniques).

Finally, we want to use a saddlepoint expansion to approximate the upper-tailed probability $P_G(M_{\underline{x}Y} \geq U_{mn})$. Jensen (1995) defined an upper-tailed probability, denoted by $P(\overline{X} \geq x)$, in equation (3.3.17) on page 79 by

$$P(\overline{X} \geq x) = (1 - \Phi(r))(1 + O(n^{-2})) + \phi(r)\left(\frac{1}{\lambda} - \frac{1}{r} + O(n^{-3/2})\right) \tag{6.14}$$

with

$$\lambda = \sqrt{n}\left(1 - e^{(-\hat{\theta}(x))}\right)\sigma(\hat{\theta}(x)) \text{ and } r = (\text{sgn}(\hat{\theta}(x)))\left(2n(\hat{\theta}(x)x - \kappa(\hat{\theta}(x)))\right)^{1/2} \tag{6.15}$$

where $\hat{\theta}(x)$ denotes the saddlepoint, $\text{sgn}(\hat{\theta}(x)) = 1, -1$ or 0 depending on whether $\hat{\theta}(x)$ is positive, negative or zero and $O(\cdot)$ is the big O function. In general, the notation $f(n) = O(g(n))$ means there is a real constant $c > 0$ and an integer n_0 such that $|f(n)| \leq c|g(n)|$ for all $n \geq n_0$ and where $f(n)$ and $g(n)$ are functions of the variable n. In other words, the notation $f(n) = O(g(n))$ states that the function $|f(n)|$ is bounded above by a constant multiple of the function $|g(n)|$ for all sufficiently large values of n indicated by $n \geq n_0$. Getting back to equations (6.14) and (6.15) it should be noted that the derivation of r was done separately on page 75 of Jensen (1995) using equations (3.3.2) and (3.3.3). The Lugannani and Rice (1980) paper was the first to give formula (6.14). Although they were the first to give formula (6.14), their paper is perhaps not easy to read. However, Daniels (1987) has given a very readable account where formula (6.14) is also given. In this thesis we mostly refer to Jensen (1995), because Jensen's textbook gives a rigorous account of the underlying mathematical theory of saddlepoint methods.

Using (6.14) and (6.15) we obtain

$$P\left(M_{\underline{x}Y} \geq U_{mn}\right) = P\left(\overline{M}_{\underline{x}Y} \geq \frac{U_{mn}}{n}\right) = P\left(\overline{M}_{\underline{x}Y} \geq \mu\right) \approx 1 - \Phi(r) + \phi(r)\left(\frac{1}{\lambda} - \frac{1}{r}\right) \quad (6.16)$$

with

$$\lambda = \sqrt{n}(1 - e^{-\gamma})\sigma(\gamma) \text{ and } r = \left(\text{sgn}(\gamma)\right)\left(2n(\gamma\mu - \kappa(\gamma))\right)^{1/2} \quad (6.17)$$

where γ denotes the saddlepoint.

Using (6.16) we can approximate the signal probability $p_G(\underline{x})$ given in Result 6.1.

6.1.6. Monte Carlo simulation

We run into a problem when computing the out-of-control and in-control unconditional average run lengths, since both formulas (see equations (6.3) and (6.7)) are m-dimensional integrals. A solution to this problem is using Monte Carlo simulation. Monte Carlo methods are based on the use of random numbers and probability statistics to investigate problems. It consists of a collection of ways for generating random samples on a computer and then using them to solve problems by providing approximate solutions to those problems. Moreover, Monte Carlo methods are useful for obtaining numerical solutions to problems which are too complicated to solve analytically and are, in this thesis, used to evaluate multiple integrals. Monte Carlo simulation is applied here to approximate the unconditional *ARL*'s for the in-control and out-of-control situations, respectively. It should be noted that these are approximations to *m*-dimensional integrals (see equations (6.3) and (6.7) for the out-of-control and in-control unconditional average run length formulas, respectively). The Monte Carlo estimates are given by (6.8) and (6.9), respectively, and by studying these formulas we see that the computations of $p_G(\underline{x})$ (for the out-of-control situation) and $p_U(u)$ (for the in-control situation) are repeated K times to obtain the Monte Carlo estimates given in (6.8) and (6.9).

Monte Carlo simulation used to approximate the unconditional ARL for the in-control situation

Chakraborti and Van de Wiel (2003) proposed five methods for computing (or approximating) ARL_0. The first three methods are similar in the sense that they all make use of Monte Carlo simulation using (6.9), but they differ in the way that $p_U(\underline{u})$ is computed or approximated. The five methods are as follows:

(i) Exact

Monte Carlo simulation is applied to approximate the ARL_0 using (6.9), with $p_U(\underline{u})$ computed exactly using (6.12).

(ii) Lugannani-Rice formula

Monte Carlo simulation is applied to approximate the ARL_0 using (6.9), with $p_U(\underline{u})$ computed approximately using (6.16).

(iii) Normal Approximation

Monte Carlo simulation is applied to approximate the ARL_0 using (6.9), with $p_U(\underline{u})$ computed using a normal approximation.

The first three methods have the same problem, namely, that we need to compute $p_U(\underline{u})$ K times for K Monte Carlo reference samples, where a reference sample is drawn from the uniform(0,1) distribution. Each element is taken from the uniform(0,1) distribution, since we're approximating the *in-control* average run length. The number of Monte Carlo reference samples K should be taken large enough so that the Monte Carlo error is acceptably small and, consequently, using methods (i), (ii) or (iii) may be time-consuming. By fixing the reference sample we would only need to compute $p_U(\underline{u})$ once. This is done in the fourth method by using the empirical cdf of $X_1, X_2, ..., X_m$.

(iv) Fixed reference sample

Recall that $F(x)$ denotes the unknown continuous cdf of each of $X_1, X_2, ..., X_m$. Let $F_m(\underline{x})$ denote the empirical cdf of $X_1, X_2, ..., X_m$. By the law of strong numbers (see, for example, Bain and Engelhardt (1992)), when m is large, the empirical cdf $F_m(\underline{x})$ converges to $F(\underline{x})$ (which is the cdf of the uniform(0,1) distribution), i.e. $F_m(\underline{x}) \to F(\underline{x})$ as $m \to \infty$, almost surely for fixed $\underline{x}$. Using this, we can replace the i^{th} reference sample observation by the $\left(i/(m+1)\right)^{th}$ quantile, $i = 1,2,...,m$, of the uniform(0,1) distribution. Since this quantile is equal to $i/(m+1)$ (say, $q_i = i/(m+1)$), we can approximate ARL_0 by $1/p_U(\underline{q})$ where $\underline{q} = (q_1, q_2, ..., q_m) = \left(1/(m+1), ..., m/(m+1)\right)$. It should be noted that one should only use the empirical cdf (and as a result fix the reference sample to $\underline{x} = \underline{u} = \underline{q}$) when m is large. Using this method we only require one reference sample and we only compute $p_U(\underline{u})$ once.

(v) Reciprocal of the false alarm rate

A quick way to approximate the ARL_0 is by using the fact that if the charting statistics, $M_{XY}^1, M_{XY}^2, ...$, were independent, the ARL_0 would be equal to the reciprocal of the false alarm rate, i.e. $ARL_0 = \frac{1}{FAR} = \frac{1}{2P\left(M_{XY} \geq U_{mn}\right)}$. When implementing this method, the FAR is estimated using the Fix-Hodges approximation formula (see Fix and Hodges (1955)). This approximation improves the normal approximation by including moments of order three and higher. Since the charting statistics are in fact dependent, we can only use the reciprocal of the false alarm rate as a quick approximation to the ARL_0. Further motivation for using the reciprocal of the false alarm rate as a quick approximation to the ARL_0 is given by Chakraborti (2000). In that paper the author showed that for the Shewhart $\overline{X}$ chart, $\frac{1}{FAR}$ can be used as a

lower bound to the ARL_0. Following this, we use the reciprocal of the false alarm rate as a quick approximation to the ARL_0.

Methods (iv) and (v) have the advantage that we don't have to draw K reference samples, since we approximate ARL_0 by $1/p_U(\underline{q})$ (using method (iv)) and by $1/2P(M_{XY} \geq U_{mn})$ (using method (v)).

The five abovementioned methods show one how to calculate (or approximate) the unconditional ARL_0 corresponding to a given value of the UCL. A table containing values of the unconditional ARL_0, for various values of m and n, is provided (see Table 1, Chakraborti and Van de Wiel (2003)). The table is given on the next page for reference. K is kept constant ($K = 1000$) to obtain a fair comparison regarding the computing times. The values in Table 6.1 were computed using all five abovementioned methods. The table shows two computing times. The first computing time is the time it took a 3.2GHz Pentium PC with 512MB of internal RAM to compute the values using Mathematica 6.0. The second computing time (given in brackets) is the computing time found by Chakraborti and Van de Wiel (2003) using a 1.7GHz Pentium PC with 128MB of internal RAM.

Certain in-control average run length values (indicated by ** in Table 6.1) could not be computed within a practical time. Chakraborti and Van de Wiel (2003) determined these computing times by multiplying the computing time for K = 1 by 1 000 and, consequently, getting a computing time for K = 1 000. In this paper the same course of action was taken to estimate the computing times for K = 1 000. From Table 6.1 we see that the 3.2GHz Pentium PC with 512MB of internal RAM is at least three times faster than the 1.7GHz Pentium PC with 128MB of internal RAM. Interpreting the times in Table 6.1 we find that the exact method is exceptionally time-consuming, particularly so as m increases. Similarly, using the LR-formula is also very time-consuming, again, particularly as m increases, but it's not as severely time-consuming as the exact method. Although fast approximations are given by the normal approximation, they are inaccurate. Fast approximations are also given by the fixed-reference-sample method and the reciprocal-of-the-false-alarm-rate method.

Table 6.1[*]**.** ARL_0 approximations and computing times for various m and n values and $U_{mn} = 4344$ [†].

m	***n***	**Exact**		**Lugannani-Rice formula**		**Normal Approximation**		**Fixed reference sample**		**Reciprocal of the false alarm rate**	
		$\hat{ARL}_0$	Time (sec.)	$\hat{ARL}_0$	Time (sec.)	$\hat{ARL}_0$	Time (sec.)	$\hat{ARL}_0$	Time (sec.)	$\hat{ARL}_0$	Time (sec.)
50	5	486	18 (54)	506	12 (36)	307	0.33 (1.00)	403	0.02 (0.05)	247	0.003 (0.01)
	10	504	132 (395)	505	12 (34)	327	0.33 (1.00)	524	0.02 (0.05)	226	0.003 (0.01)
	25	488	1618 (4850)	491	10 (31)	425	0.40 (1.20)	694	0.02 (0.05)	119	0.003 (0.01)
100	5	496	75 (220)	505	18 (48)	219	0.40 (1.20)	478	0.02 (0.05)	353	0.003 (0.01)
	10	505	640 (1920)	506	16 (47)	339	0.42 (1.30)	531	0.02 (0.05)	332	0.003 (0.01)
	25	**	8900 (26168)	503	18 (48)	422	0.42 (1.30)	683	0.03 (0.06)	233	0.003 (0.01)
500	5	491	3544 (10633)	496	70 (207)	226	0.40 (1.20)	492	0.07 (0.20)	445	0.003 (0.01)
	10	**	24500 (73516)	513	60 (179)	367	0.60 (1.70)	537	0.07 (0.21)	484	0.003 (0.01)
	25	**	$2.53*10^5$ ($7.59*10^5$)	494	60 (176)	445	0.55 (1.60)	578	0.10 (0.29)	450	0.003 (0.01)
1000	5	**	10601 (31766)	500	120 (356)	235	0.70 (2.10)	513	0.16 (0.48)	471	0.003 (0.01)
	10	**	$1.15*10^5$ ($3.42*10^5$)	499	126 (373)	355	0.80 (2.40)	516	0.18 (0.49)	488	0.003 (0.01)
	25	**	$1.05*10^6$ ($3.15*10^6$)	500	117 (348)	442	0.61 (1.70)	548	0.20 (0.63)	482	0.003 (0.01)
2000	5	**	$0.57*10^5$ ($1.71*10^5$)	503	240 (713)	234	0.70 (2.10)	506	0.22 (0.67)	474	0.003 (0.01)
	10	**	$0.48*10^6$ ($1.44*10^6$)	504	221 (659)	354	0.64 (1.90)	513	0.24 (0.71)	499	0.003 (0.01)
	25	**	$0.43*10^7$ ($1.29*10^7$)	509	229 (676)	446	0.71 (2.10)	531	0.48 (1.41)	497	0.003 (0.01)

[*] Chakraborti and Van de Wiel (2003) wrote a Mathmatica program to approximate the ARL_0 for a given m, n and value of the UCL. This Mathematica program can be downloaded using the website www.win.tue.nl/~markvdw. For more details on this Mathematica program see Mathematica Program 1 in Appendix B.

[†] Table 6.1 appears in Chakraborti and Van de Wiel (2003), Table 1. It should be noted that Chakraborti and Van de Wiel (2003) failed to say what the value of the UCL was set equal to when constructing this table. Turning to their Mathematica program we see that the user specific parameters are set equal to m=1000, n=5 and UCL=4344. Recall that only the UCL needs to be specified, since the LCL can be calculated using $L_{mn}=mn-U_{mn}$.

In Table 6.1 we have the exact formula and various approximations for the unconditional ARL_0. We see that although the exact formula gives us ARL_0 values close to 500 (which is desirable), the exact computations are very time-consuming for most values of m and n. Focusing on the approximations, the closer an ARL_0 value is to 500, the better the approximation. Using this criteria we see that the normal approximation is inaccurate for all values of m and n. The fixed-reference-sample and the reciprocal-of-the-false-alarm-rate approximations are relatively good for $m \geq 1000$ and, in particular, the fixed-reference-sample approximation performs better for 'small' values of n ($=5$ or 10) than for n 'large' ($n = 25$). It seems that the best approximation is the LR-formula, since all the corresponding ARL_0 values are close to 500. In summary, the best method of calculating the unconditional ARL_0 is by using the exact formula, if it's not too time-consuming, otherwise the LR-formula is the best approximation.

Monte Carlo simulation used to approximate the unconditional ARL for the out-of-control situation

Monte Carlo simulation is used to approximate the unconditional ARL for the out-of-control situation. There are concerns about the number of Monte Carlo samples used, namely, that although larger sizes of K will result in more accurate approximations and smaller Monte Carlo errors, using larger Monte Carlo samples may be time-consuming or computationally expensive or both.

Since the unconditional ARL is the average of the conditional $ARL_G(\underline{X})$ over all possible $\underline{X}'s$ and the K Monte Carlo reference samples are independent, the Monte Carlo standard error of the estimate $\hat{ARL}$ is given by

$$\sigma_{mc} = \frac{\sigma\left(ARL_G(\underline{X})\right)}{\sqrt{K}} \tag{6.18}$$

where $\sigma\left(ARL_G(\underline{X})\right)$ denotes the unknown standard deviation of $ARL_G(\underline{X})$. From (6.18) we see that the standard error decreases with the square root of the number of Monte Carlo samples used. If we, for example, quadruple the number of Monte Carlo samples used, we will half the standard error. While increasing K is one technique for reducing the standard

error, doing so can be time-consuming or computationally expensive or both. Clearly, a better approach is needed.

Let D denote some specified value such that

$$s_{mc} = \frac{s\left(ARL_G(\underline{X})\right)}{\sqrt{K}} \leq D. \tag{6.19}$$

where the sample standard deviation $s\left(ARL_G(\underline{X})\right)$ is used to estimate $\sigma\left(ARL_G(\underline{X})\right)$ and, subsequently, $s_{mc} = \frac{s\left(ARL_G(\underline{X})\right)}{\sqrt{K}}$ is used to estimate $\sigma_{mc} = \frac{\sigma\left(ARL_G(\underline{X})\right)}{\sqrt{K}}$.

We want to find the smallest K such that (6.19) is satisfied. We start by taking K 'small', say $K = 100$, for example, and then we compute the corresponding standard error s_{mc}. If (6.19) is not satisfied we increase K and the process is repeated until the standard error is smaller than or equal to some specified value D. It should be noted that D could also be taken to be some percentage of the estimate $A\hat{R}L$. By implementing (6.19), we find an accurate approximation (with a small Monte Carlo error) of the unconditional ARL for the out-of-control situation.

6.1.7. Determination of chart constants

Up to this point we've addressed the problem where one has to calculate the (unknown) unconditional ARL_0 for a given (known) upper control limit. In this section we address the opposite problem where one has to calculate the (unknown) upper control limit for a specified (known) ARL_0. In order to solve the latter problem, we use an iterative procedure based on linear interpolation. An initial value for the *UCL*, say $UCL^{(1)}$, is needed to start the iteration. We can limit our search of $UCL^{(1)}$ (and ultimately of *UCL*) to integer values between 0 and *mn*, since the MW charting statistic only takes on integer values between 0 and *mn* (recall that $0 \leq M_{XY} \leq mn$). In addition, we use the fact that ARL_0 is strictly increasing in *UCL* (and subsequently, $p_U(\underline{u})$ is strictly decreasing in *UCL*). Let the desired unconditional $ARL_0 = 500$.

To obtain $UCL^{(1)}$, the fixed-reference-sample approximation or the reciprocal-of-the-FAR approximation can be used, since they are both very fast and relatively accurate approximations. For the latter approach, we equate the reciprocal of the false alarm rate to 500, meaning that we have to solve for $\frac{1}{2(FH(u))} = 500$, where $FH(u)$ denotes the Fix-Hodges approximation for the upper tail probability $P_0(M_{\underline{xY}} \geq u)$. We estimate (through Monte Carlo simulation) ARL_0 at $UCL^{(1)}$ using (6.9), where the LR-approximation is used to calculate $p_U(\underline{u})$. In doing so, we obtain a new ARL_0, say $ARL_0^{(1)}$. If $ARL_0^{(1)}$ is smaller than 500, we increase the value of $UCL^{(1)}$ by a specified amount, say s, to obtain $UCL^{(2)} = UCL^{(1)} + s$. On the contrary, if $ARL_0^{(1)}$ is greater than 500, we decrease the value of $UCL^{(1)}$ to obtain $UCL^{(2)} = UCL^{(1)} - s$. Using $UCL^{(2)}$, the search procedure is repeated until ARL_0 is 'satisfactorily close' to the target value of 500.

A question arises: How close is 'satisfactorily close'? To answer this Chakraborti and Van de Wiel (2003) suggest using a target interval, say $500 \pm \lambda 500$, where λ denotes the percentage deviation from the target value that is acceptable. Suppose we allow a deviation of 3%, i.e. $\lambda = 0.03$, the search procedure stops at the l^{th} step if $485 \leq ARL_0^{(l)} \leq 515$. The larger this margin, the faster the algorithm, and as a result, the faster a solution is found. If the specifications can't be met, the algorithm returns one or more solutions for which ARL_0 is close to the target value. If $\lambda = 0$, the search procedure stops at the l^{th} step if $UCL^{(l)}$ has a corresponding $ARL_0 < 500$ and $UCL^{(l)} + 1$ has a corresponding $ARL_0 \geq 500$, and as a result, the practitioner has to decide whether to use $UCL^{(l)}$ or $UCL^{(l)} + 1$.

We illustrate this search procedure with an example. Suppose the reference sample size is 50 (m = 50), the test sample size is 5 (n = 5) and we want to find the chart constants (U_{mn} and $L_{mn} = mn - U_{mn}$) such that the specified target $ARL_0 = 500$. Suppose we specify that a 3% deviation from the target value is acceptable, i.e. $\lambda = 0.03$. In doing so, the search procedure stops when $485 \leq \hat{ARL}_0 \leq 515$ and yields the corresponding chart constants. In addition, we specify that the Monte Carlo standard error, s_{mc}, be smaller than or equal to 2.5% of the estimate of ARL_0. Then $D = 0.025 \times 500 = 12.5$ is the maximum value of the

standard error of the estimate that we allow. The output from the program is given in Table 6.2.

Table 6.2. Finding chart constants for m=50, n=5, target ARL_0=500, $\lambda = 0.03$ and $D = 12.5$*.

1/(false alarm rate approximation)
(1) ucl= 222 lcl= 28 ARL0= 500
Fixed reference sample approximation
(2) ucl= 222 lcl= 28 ARL0= 874.22
(3) ucl= 212 lcl= 38 ARL0= 206.763
(4) ucl= 216 lcl= 34 ARL0= 351.068
(5) ucl= 218 lcl= 32 ARL0= 467.529
LR-approximation
(6) ucl= 218 lcl= 32 ARL0= 571.016 smc= 15.4659 5% perc= 122.654 K= 2000
(7) ucl= 208 lcl= 42 ARL0= 138.825 smc= 3.46778 5% perc= 44.3363 K= 1061
(8) ucl= 216 lcl= 34 ARL0= 426.735 smc= 11.7639 5% perc= 95.169 K= 2000
(9) ucl= 217 lcl= 33 ARL0= 494.728 smc= 13.6655 5% perc= 105.761 K= 2000
{231.156,Null}

From Table 6.2 it can be seen that one iteration has been carried out under the reciprocal-of-the-*FAR* approximation and that four iterations have been carried out under the fixed-reference-sample approximation. These five iterations didn't take long, since both the reciprocal-of-the-*FAR* and the fixed-reference-sample approximations are fast approximations. For each of these five iterations, the values of U_{mn} (denoted ucl in the output), L_{mn} (denoted lcl in the output) and the corresponding unconditional ARL_0 (denoted ARL0 in the output) are given.

From Table 6.2 it can also be seen that four iterations have been carried out under the Lugannani-Rice approximation. For each of these four iterations, the values of U_{mn}, L_{mn}, the corresponding unconditional ARL_0, the standard error of the estimated ARL_0 (denoted smc in the output), the estimated 5th percentile of the conditional in-control *ARL* distribution (denoted 5% perc in the output) and the number of Monte Carlo samples used to obtain the estimates (denoted by K in the output) are given. In total there were nine iterations that have been carried out in approximately 231 seconds. The final chart constants are found at iteration number 9. They are $U_{mn} = 217$ and $L_{mn} = 33$ with a corresponding unconditional

* The values in Table 6.2 were obtained by running the Mathematica program provided by Chakraborti and Van de Wiel (2003). See Mathematica Program 1 in Appendix B for more information on this Mathematica program.

$ARL_0 = 494.728$. The 5th percentile of the conditional in-control *ARL* distribution is equal to 105.761, meaning that 95% of all reference samples (that could possibly have been taken from the in-control process) will generate a conditional ARL_0 of *at least* 106. Previously stated is the fact that K is chosen such that the standard error of the estimate is smaller than or equal to 2.5% of the estimate. Stated differently, K is chosen such that $s_{mc} \leq D$. When studying iteration number 9, we see that this condition is not satisfied, since $s_{mc} = 13.6655 > 12.5$. The reason for this is that the maximum number of Monte Carlo samples is set to 2 000 in the Mathematica program. If there were no restriction put on the number of Monte Carlo samples, the iterative procedure would have increased K and repeated the process until the standard error is smaller than or equal to $D = 12.5$. Table 6.3 contains the chart constants for various values of *m* and *n* with $\lambda = 0.03$ and where s_{mc} must be smaller than or equal to 2.5% of the estimate of ARL_0.

Table 6.3. Control limits for various values of *m* and *n*[*].

m	***n***	$ARL_0 = 370$		$ARL_0 = 500$	
		L_{mn}	U_{mn}	L_{mn}	U_{mn}
50	5	35	215	33	217
	10	115	385	111	389
	25	400	850	393	857
100	5	69	431	65	435
	10	231	769	224	776
	25	805	1695	793	1707
500	5	348	2152	328	2172
	10	1170	3830	1128	3872
	25	4081	8419	4016	8484
1000	5	698	4302	653	4347
	10	2344	7656	2268	7732
	25	8169	16831	8058	16942
2000	5	1397	8603	1309	8691
	10	4682	15318	4540	15460
	25	16392	33608	16145	33855

[*] The values in Table 6.3 were obtained by running the Mathematica program provided by Chakraborti and Van de Wiel (2003). See Mathematica Program 1 in Appendix B for more information on this Mathematica program. Table 6.3 also appears in Chakraborti and Van de Wiel (2003), Table 3.

Example 6.1

A Mann-Whitney control chart based on the Montgomery (2001) piston ring data

For the piston-ring data with $m = 125$ (see example 4.1 for an explanation of why m is equal to 125 (and not 25 like some of the earlier examples)) and $n = 5$, Chakraborti and Van de Wiel (2003) found the upper and lower control limits of the Shewhart-type MW chart to be 540 and 85, respectively. The control limits are obtained by setting the user-specific parameters equal to $m = 125$, $n = 5$, target $ARL_0 = 400$, $\lambda = 0.02$ and $D = 6$ in the Mathematica program provided by Chakraborti and Van de Wiel (2003). By setting $D = 6$ we require that $s_{mc} \leq 6$. By setting $\lambda = 0.02$ we specify that a 2% deviation from the target value is acceptable. In doing so, the search procedure stops when $392 \leq \hat{ARL}_0 \leq 408$ and yields the corresponding chart constants, which (in this case) equal $L_{mn} = 85$ and $U_{mn} = 540$. The fifteen Phase II samples and the reference sample lead to fifteen MW statistics shown in Table 6.4 (read from left to right and to left) and the MW control chart is shown in Figure 6.1.

Table 6.4. Phase II MW statistics for the Piston-ring data in Montgomery (2001)*.

429.0	333.0	142.5	370.5	241.5	410.5	393.0	240.5
471.0	486.0	340.5	561.0	575.5	601.5	484.5	

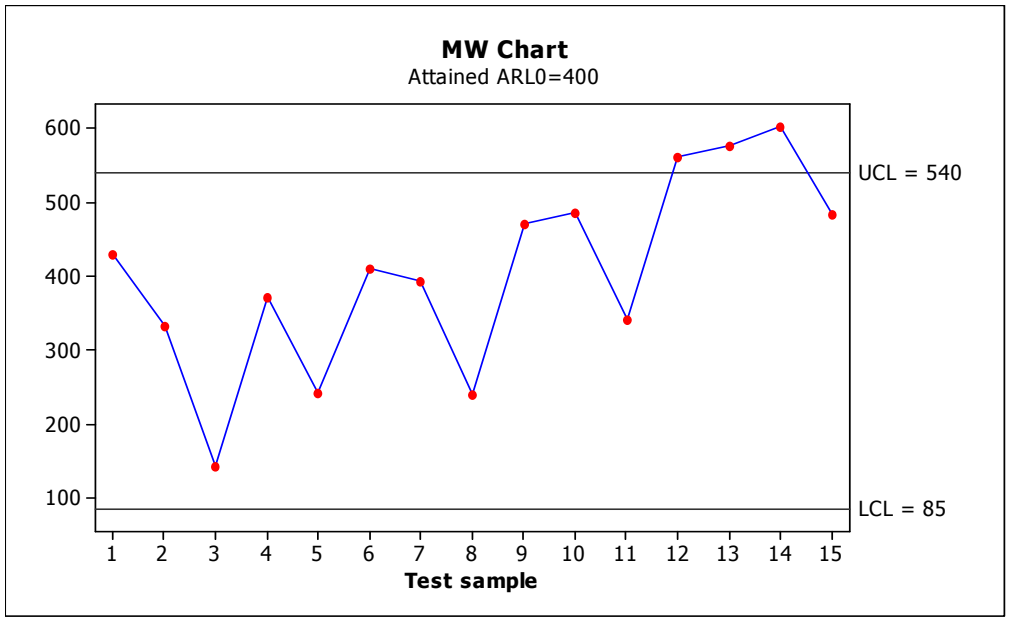

Figure 6.1 MW Chart for the Montgomery (2001) piston ring data.

*The values in Table 6.4 were calculated using Minitab.

It is seen that all but three of the test groups, 12, 13 and 14 are in-control. The conclusion from the MW chart is that the medians of test groups 12, 13 and 14 have shifted to the right in comparison with the median of the in-control distribution, assuming that G is a location shift of F. It may be noted that the Shewhart $\overline{X}$ chart shown in Montgomery (2001) led to the same conclusion with respect to the means. Of course, the advantage with the MW chart (and with any nonparametric chart) is that it is distribution-free, so that regardless of the underlying distribution, the in-control ARL of the chart is roughly equal to 400 and there is no need to worry about (non-) normality, as one must for the $\overline{X}$ chart. For comparison purposes, the distribution-free *1-of-1* precedence chart for this data for an unconditional $ARL_0 = 400$ is found to be $LCL = 73.982$ and $UCL = 74.017$ for an attained $ARL_0 \approx 414.0$. Consequently, the precedence chart declares the 12^{th} and the 14^{th} groups to be out of control but not the 13^{th} group, unlike the MW and the Shewhart chart. This is not entirely surprising since the MW test is generally more powerful than the precedence test.

6.1.8. Control chart performance

The performance of a control chart is usually judged in terms of certain characteristics associated with its run-length distribution. For the most part the ARL is used to evaluate chart performance, since it indicates, on average, how long one has to wait before the chart signals. Some researchers have advocated using other characteristics than the ARL, such as percentiles of the run length distribution (see Section 2.1.5 for a detailed discussion on this issue). Chakraborti and Van de Wiel (2003) examined the ARL, the 5^{th} and the 95^{th} percentiles (denoted ρ_5 and ρ_{95}, respectively) of the *conditional* distribution for the MW chart. The question may be raised about why the authors decided to use (only) the *conditional* distribution when both the conditional *and* unconditional distributions provide key information concerning the performance of a chart. Recall that the unconditional distribution results from averaging over all possible reference samples and in practice researchers would (almost certainly) not have the benefit of averaging.

Chakraborti and Van de Wiel (2003) compared the MW chart to the Shewhart $\overline{X}$ chart. For the latter we assume case UU, when both the mean and variance are unknown, and consequently both parameters need to be estimated from the reference sample. Therefore, the MW chart is compared to the Shewhart $\overline{X}$ chart *with estimated parameters*. Additionally,

both charts are designed to have the same in-control average run length ($ARL_0 \approx 500$). The latter two conditions are necessary to ensure a fair comparison between the MW and Shewhart $\overline{X}$ chart. The in-control case is considered first.

In-control performance

For the in-control case a lower order percentile, specifically the 5th percentile, should be examined (the 95th percentile is also examined for completeness). Recall that we want the in-control ARL to be large and, by the same token, large values of the 5th percentile are desirable. The test sample size, n, was taken to equal 5 for all cases, whereas the reference sample size, m, varies from 50 to 2000. Both the reference and test samples were drawn from a normal distribution, specifically a $N(0,1)$ distribution. The results were obtained using $K = 1000$ simulations and are shown in Table 6.5.

Table 6.5. The 5th and 95th percentiles and standard deviations of the conditional in-control distribution with $n = 5$ and $ARL_0 \approx 500$ *.

m	MW chart				Shewhart $\overline{X}$ Chart			
	Upper Control Limit	ρ_5	ρ_{95}	Standard Deviation	Upper Control Limit	ρ_5	ρ_{95}	Standard Deviation
50	217	97	1292	553	3.01996	49	1619	854
75	326	146	1219	461	3.05156	87	1379	645
100	435	182	1146	358	3.06535	112	1290	463
150	654	251	1090	315	3.07715	154	1197	377
300	1304	284	845	197	3.08607	232	927	235
500	2172	322	700	140	3.08848	270	828	174
750	3258	360	677	107	3.08935	314	765	140
1000	4347	379	674	83	3.08969	338	721	121
2000	8691	420	629	55	3.09007	376	651	84

From Table 6.5 we find that for the MW chart with $m = 100$ and a control chart constant of 435 ($U_{mn} = 435$), $\rho_5 = 182$, meaning that 95% of the in-control average run lengths are at least 182, whereas for the Shewhart $\overline{X}$ chart with $m = 100$ and a control chart

* The values in Table 6.5 were obtained by running the Mathematica program provided by Chakraborti and Van de Wiel (2003). See Mathematica Program 1 in Appendix B for more information on this Mathematica program. Table 6.5 also appears in Chakraborti and Van de Wiel (2003), Table 4.

constant of 3.06535, $\rho_5 = 112$, meaning that 95% of the in-control average run lengths are at least 112. Since 182 > 112 it can be concluded that the in-control performance of the MW chart is better than that of the Shewhart $\overline{X}$ chart with estimated parameters. Moreover, all the 5th percentiles of the MW chart are larger than those of the Shewhart $\overline{X}$ chart with estimated parameters, particularly for $m \leq 150$, further supporting the statement that the in-control performance of the MW chart is better than that of the Shewhart $\overline{X}$ chart with estimated parameters. The estimated standard deviations are given in Table 6.5 to give some information about the variability of ARL_0. All the estimated standard deviations for the MW chart are smaller than those of the Shewhart $\overline{X}$ chart with estimated parameters, further supporting the statement that the in-control performance of the MW chart is better.

Out-of-control performance

For the out-of-control case a higher order percentile, specifically the 95th percentile, is examined. Recall that we want the out-of-control ARL to be small and, by the same token, small values of the 95th percentile are desirable. These two performance measures, the ARL_δ and ρ_{95}, are examined for three distributions, namely, the Normal, Laplace and Gamma(2,2) distributions, respectively. The motivation for examining these three distributions is that we would like to examine a symmetric (Normal), asymmetric (Gamma(2,2)) and heavy-tailed (Laplace) distribution, respectively. The Laplace distribution is comparable to the Normal distribution, but it has heavier tails, while the Gamma(2,2) distribution is positively skewed.

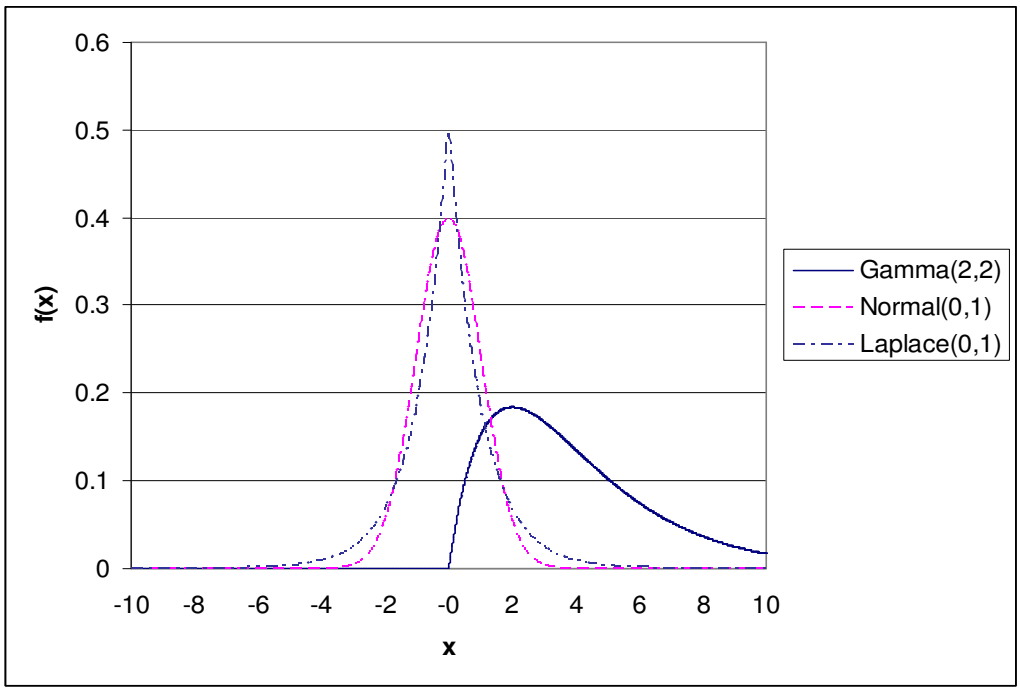

Figure 6.2. The shapes of the three distributions under consideration.

(i) The normal distribution

For the Normal distribution we would expect the out-of-control performance of the Shewhart $\overline{X}$ chart with estimated parameters to be better than that of the MW chart. The reason for this being that it's typical for normal theory methods to outperform nonparametric methods when the normality assumption is met. A two-sided chart was applied in the case of the Normal distribution. The test sample size, n, was taken to equal 5, whereas the reference sample size, m, was taken to equal 100. The chart constants for both the MW and Shewhart $\overline{X}$ chart are chosen such that the in-control average run length is approximately equal ($ARL_0 \approx 500$) for both charts. ARL_δ and the 95th percentiles of the distribution of ARL_δ were computed, using these chart constants, for various values of δ, where δ is the unknown shift parameter (recall that shift alternatives are denoted as $G(x) = F(x-\delta)$). The results are shown below in Figure 6.3.

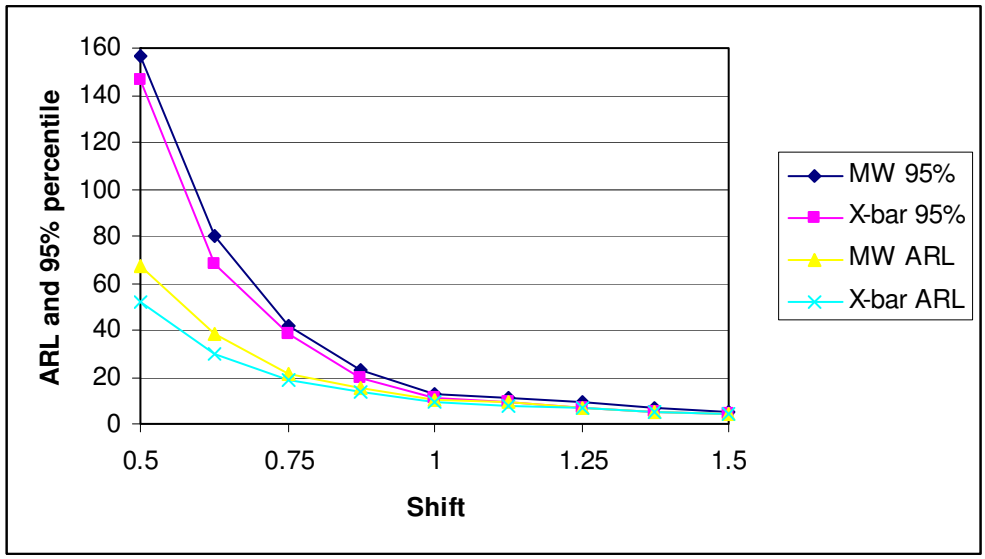

Figure 6.3. Comparison of the MW chart with the Shewhart $\overline{X}$ chart for the Normal distribution.

When comparing the MW chart with the Shewhart $\overline{X}$ chart under Normal shift alternatives we find that the Shewhart $\overline{X}$ chart is performing only slightly better than the MW chart, since the 95th percentiles for the Shewhart $\overline{X}$ chart are smaller than the 95th percentiles for the MW chart. However, it should be noted that the differences are small and it appears to fade away when the shift is greater than one. A similar pattern holds for the ARL_δ's.

(ii) The Laplace distribution

The Laplace distribution, also called the Double-Exponential distribution, is comparable to the Normal distribution, but it has heavier tails (see Figure 6.2). As a result, there are higher probabilities associated with extreme values when working with the Laplace distribution as opposed to using the Normal distribution. For the Laplace distribution we would expect the out-of-control performance of the MW chart to be better than that of the Shewhart $\overline{X}$ chart. The reason for this being that it's typical for nonparametric methods to outperform normal theory methods when the distribution in question is heavy-tailed (see, for example, Gibbons and Chakraborti (2003)). A two-sided chart was applied to the Laplace distribution. For consistency, $n = 5$ and $m = 100$ (the same values were used under Normal shift alternatives) and the chart constants for both the MW and Shewhart $\overline{X}$ chart are chosen

such that the in-control average run length is approximately equal ($ARL_0 \approx 500$) for both charts.

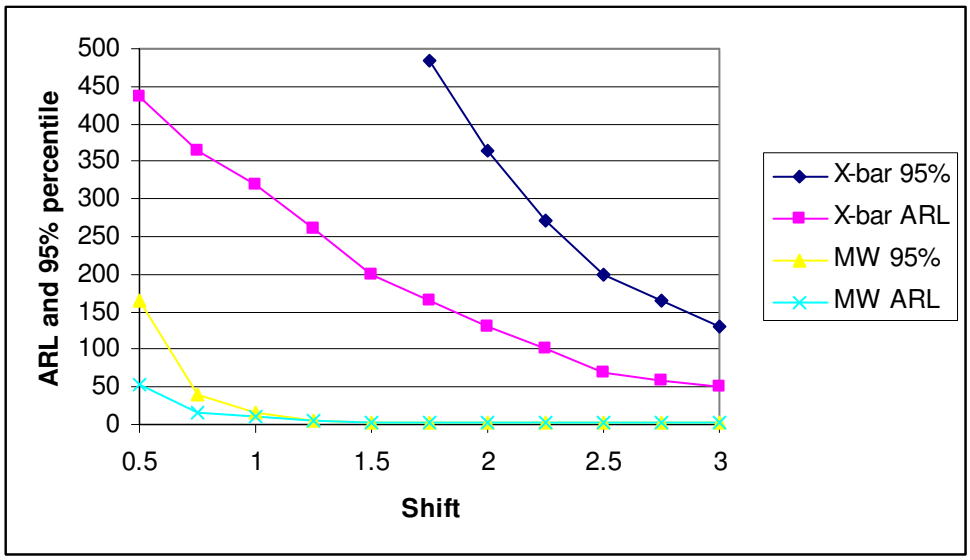

Figure 6.4. Comparison of the MW chart with the Shewhart $\overline{X}$ chart for the Laplace distribution.

When comparing the MW chart with the Shewhart $\overline{X}$ chart under Laplace shift alternatives we find that the MW chart is performing much better than the Shewhart $\overline{X}$ chart, since the 95th percentiles for the MW chart are smaller than the 95th percentiles for the Shewhart $\overline{X}$ chart. It should be noted that these differences are reasonably large for all shifts, indicating that the MW chart is performing a great deal better than the Shewhart $\overline{X}$ chart.

(iii) The Gamma distribution

From Figure 6.2 it can be seen that the Gamma(2,2) distribution is positively skewed. For the Gamma(2,2) distribution we would expect the out-of-control performance of the MW chart to be better than that of the Shewhart $\overline{X}$ chart. An upper one-sided chart was applied to the Gamma(2,2) distribution. For consistency, $n = 5$ and $m = 100$ (the same values were used under Normal and Laplace shift alternatives) and the chart constants for both the MW and Shewhart $\overline{X}$ chart are chosen such that the in-control average run length is approximately equal ($ARL_0 \approx 500$) for both charts.

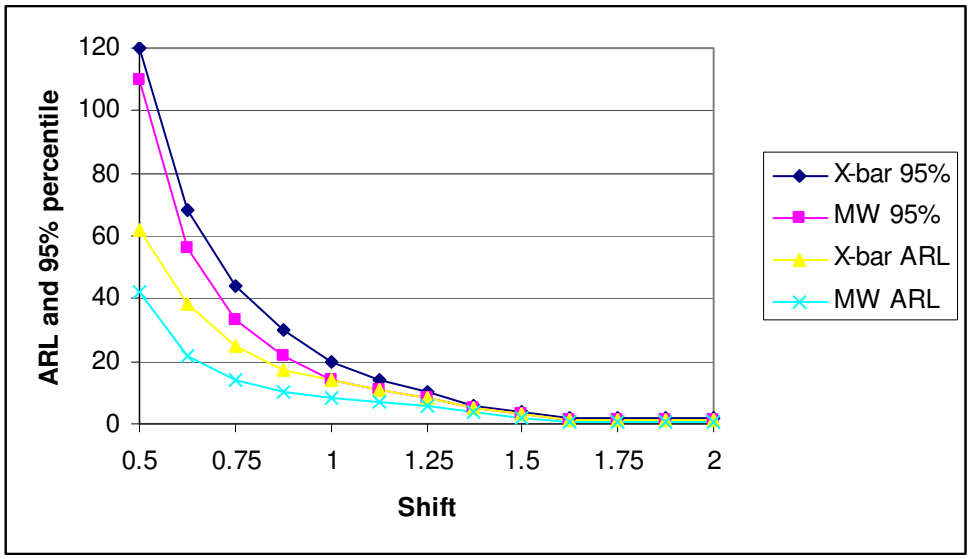

Figure 6.5. Comparison of the MW chart with the Shewhart $\overline{X}$ chart for the Gamma distribution.

When comparing the MW chart with the Shewhart $\overline{X}$ chart under Gamma(2,2) shift alternatives we find that the MW chart is performing better than the Shewhart $\overline{X}$ chart, since the 95th percentiles for the MW chart are smaller than the 95th percentiles for the Shewhart $\overline{X}$ chart. It should be noted that these differences are not as large as the differences obtained using the Laplace distribution.

The graphs were also constructed for larger values of m, but since the graphs were very similar to the given figures, they we omitted. In conclusion we found that the MW chart performs better than the Shewhart $\overline{X}$ chart with estimated parameters under heavy tailed and skewed distributions.

6.1.9. Summary

In Section 6.1 we examined a Shewhart-type chart based on the Mann-Whitney-Wilcoxon statistic. We illustrated these procedures using the piston ring data from Montgomery (2001) to help the reader to understand the subject more thoroughly. One practical advantage of the nonparametric Shewhart-type Mann-Whitney control chart is that there is no need to assume a particular parametric distribution for the underlying process (see Section 1.4 for other advantages of nonparametric charts).

6.2. The tabular Phase I CUSUM control chart

6.2.1. Introduction

Zhou, Zou and Wang (2007) (hereafter ZZW) proposed a Phase I CUSUM control chart for individual observations based on the Mann-Whitney-Wilcoxon statistic. They compared their proposed control chart to the likelihood ratio test (LRT) chart of Sullivan and Woodall (1996) and the CUSUM LT chart of Koning and Does (2000).

Suppose that a sample of size n, $X_1, X_2, ..., X_n$, is available with an unknown continuous cdf, $F(x, \mu_i)$ $i = 1,2,...,n$, where μ_i denotes the location parameter. In this set up, let μ_i denote the population mean of X_i. An out-of-control condition is a shift in the location parameter to some different value. The problem of detecting a shift in a parameter of the process is similar to sequential change-point detection (see, for example, Hawkins and Zamba (2005)). Various authors have studied the change-point problem; see for example Hawkins (1977), Sullivan and Woodall (1996), Hawkins, Qiu and Kang (2003) and Hawkins and Zamba (2005). In a change-point model, all the observations up to the change-point have the same distribution, say $F(x, \mu_a)$, while the remaining observations have the same distribution, say $F(x, \mu_b)$, i.e.

$$X_i = \begin{cases} F(x, \mu_a) & \text{for} \quad i = 1,2,...,t \\ F(x, \mu_b) & \text{for} \quad i = t+1, t+2,...,n \end{cases}$$

where t, with $1 \le t < n$, is the change-point. If $\mu_a = \mu_b$ the process is said to be in-control, whereas if $\mu_a \neq \mu_b$ the process is declared to be out-of-control. ZZW give an estimate for the position of the change-point, $\hat{\tau}$, as $\hat{\tau} = \underset{1<t<n}{\arg\max}\{|SMW_t|\}$ [*] (see Pettitt (1979)), where SMW_t is defined in (6.21). One can also look for multiple shifts, especially if the dataset is large. This could be done by partitioning the data at the location of the change-point then repeating the process on each subset of observations. This continues until no evidence of additional change-points is given. For example, if there are two shifts we have

[*] *Arg max* stands for the argument of the maximum, that is, the value of the given argument for which the value of the given expression attains its maximum value. For example, arg max $\{f(x)\}$ is the value of x for which $f(x)$ has the largest value.

$$X_i = \begin{cases} F(x,\mu_a) & \text{for} \quad i = 1,2,...,\tau_1 \\ F(x,\mu_b) & \text{for} \quad i = \tau_1+1,\tau_1+2,...,\tau_2 \\ F(x,\mu_c) & \text{for} \quad i = \tau_2+1,\tau_2+2,...,n \end{cases}$$

where τ_1 and τ_2 are the two change-points, respectively. ZZW lay emphasis on the fact that their proposed CUSUM Mann-Whitney chart is not intended to be used for detecting multiple shifts, but they still expect the chart to have good detecting performance if the mean shifts (μ_a, μ_b and μ_c) are all in the same direction, i.e. μ_a, μ_b and μ_c form either a decreasing or an increasing sequence.

The Mann-Whitney statistic* is defined to be the number of (X_i, X_j) pairs with $X_i > X_j$ where $i = 1,2,...,t$ and $j = t+1, t+2,...,n$. This can be written as

$$MW_t = \sum_{i=1}^{t}\sum_{j=t+1}^{n} I(X_j < X_i) \text{ for } t = 1,2,...,n-1 \qquad (6.20)$$

where $I(X_j < X_i)$ is the indicator function, i.e.

$$I(X_j < X_i) = \begin{cases} 1 & \text{if} \quad X_j < X_i \\ 0 & \text{if} \quad X_j \geq X_i \end{cases}.$$

The expected value, variance and standard deviation of the Mann-Whitney statistic is easy to find by using the relationship

$$MW_t = W_t - \frac{t(t+1)}{2}$$

where W_t is the well-known Wilcoxon rank-sum test statistic, that is, $W_t = \sum_{i=1}^{t} R_i$ and $R_1, R_2,..., R_t$ are the ranks of the t observations $x_1, x_2,..., x_t$ in the complete sample of n observations. The expected value and variance of W_t is given by (see Gibbons and

* The MW_t statistic is directly related to the well-known Mann-Whitney U test statistic (see Gibbons and Chakraborti (2003)) where the Mann-Whitney U test statistic is defined as the number of times Y precedes X in the combined ordered arrangement of the two samples, $X_1, X_2,..., X_m$ and $Y_1, Y_2,..., Y_n$, into a single sequence of $N = m+n$ variables. Then the U test statistic is defined as $U = \sum_{i=1}^{m}\sum_{j=1}^{n} D_{ij}$ where $D_{ij} = 1, 0$ if $Y_j < X_i$, $Y_j > X_i$ $\forall i, j$.

Chakraborti (2003)) $E(W_t)=\frac{t(n+1)}{2}$ and $\text{var}(W_t)=\frac{t(n-t)(n+1)}{12}$. As a result, the expected value, variance and standard deviation of MW_t is given by

$$E(MW_t)=E\left(W_t-\frac{t(t+1)}{2}\right)=E(W_t)-\frac{t(t+1)}{2}=\frac{t(n-t)}{2},$$

$$\text{var}(MW_t)=\text{var}\left(W_t-\frac{t(t+1)}{2}\right)=\frac{t(n-t)(n+1)}{12}, \text{ and}$$

$$stdev(MW_t)=\sqrt{\text{var}(MW_t)}=\sqrt{\frac{t(n-t)(n+1)}{12}}.$$

It follows that the standardized value of MW_t is given by

$$SMW_t=\frac{MW_t-E(MW_t)}{stdev(MW_t)}=\frac{MW_t-\frac{t(n-t)}{2}}{\sqrt{\frac{t(n-t)(n+1)}{12}}}. \qquad (6.21)$$

If all X_i-observations $(i=1,2,...,t)$ are smaller than the X_j-observations $(j=t+1, t+2,...,n)$, MW_t would be equal to zero. On the other hand, if all X_i-observations are greater then the X_j-observations, MW_t would equal $t\times(n-t)$. Therefore we have that $0\le MW_t\le t(n-t)$.

If the process is in-control, the distribution of MW_t is symmetric about its mean, $\frac{t(n-t)}{2}$ for each *t*, and large values of MW_t, that is, if a large number of X_i-observations are greater than the X_j-observations, would be indicative of a negative shift, whereas small values of MW_t would be indicative of a positive shift. If there are ties present, i.e. if any $X_i=X_j$, then recall that for a continuous random variable the probability of any particular value is zero; thus, $P(X=a)=0$ for any a. Since the distribution of the observations is assumed to be continuous, $P(X_i-X_j=0)=0$. Theoretically, ties should occur with zero probability, but in practice ties do occur. In case of the occurrence of ties, a correction to the variance of MW_t can be made by multiplying the variance by the factor $1-\frac{\sum_{i=1}^{r}g_i(g_i^2-1)}{n(n^2-1)}$ where g_i denotes the frequency of the i^{th} value and *r* denotes the distinct number of values in

the total of n observations, respectively. Since the sum over all frequencies equal n we have that $\sum_{i=1}^{r} g_i = n$. Consequently, the variance of MW_t (which is also the variance of W_t) is given by

$$\text{var}(MW_t) = \text{var}(W_t) = \left(\frac{t(n-t)(n+1)}{12}\right)\left(1 - \frac{\sum_{i=1}^{r} g_i(g_i^2 - 1)}{n(n^2-1)}\right).$$

The charting statistic for the proposed CUSUM Mann-Whitney chart is obtained by replacing y_i by SMW_t in the equations for the classic standardized CUSUM chart (these equations are given in Section 2.3 numbers (2.35), (2.36) and (2.37)).

The resulting upper one-sided CUSUM is given by

$$S_i^+ = \max[0, S_{i-1}^+ + SMW_i - k] \quad \text{for } i = 1,2,3,... \tag{6.22}$$

while the resulting lower one-sided CUSUM is given by

$$S_i^- = \min[0, S_{i-1}^- + SMW_i + k] \quad \text{for } i = 1,2,3,... \tag{6.23}$$

or

$$S_i^{-*} = \max[0, S_{i-1}^{-*} - SMW_i - k] \quad \text{for } i = 1,2,3,... \tag{6.24}$$

The two-sided CUSUM is constructed by running the upper and lower one-sided CUSUM charts simultaneously and signals at the first i such that $S_i^+ \geq h$ or $S_i^- \leq -h$. The starting values, S_0^- and S_0^+, are typically set equal to zero, that is, $S_0^- = 0$ and $S_0^+ = 0$.

6.2.2. Determination of chart constants

ZZW take the reference value, k, to equal 2. They motivate their choice of k by stating that smaller values of k lead to quicker detection of smaller shifts. Their simulation studies also support the decision of setting $k = 2$, since the simulation results show that the corresponding control chart has good performance. The decision interval, h, is chosen such that a desired *FAP*, denoted by α, is attained. ZZW considered h for various combinations of α and n. The table is given below for reference.

Table 6.6. Simulated h values for the CUSUM Mann-Whitney chart[*].

	α							
n	**0.07**	**0.06**	**0.05**	**0.04**	**0.03**	**0.02**	**0.01**	**0.0075**
20	0.753	0.885	1.053	1.306	1.656	2.194	3.247	3.671
30	1.267	1.490	1.788	2.187	2.719	3.615	5.531	6.371
40	1.774	2.111	2.525	3.081	3.882	5.258	7.612	9.134
50	2.362	2.779	3.329	4.102	5.194	6.988	10.236	11.993
60	2.940	3.454	4.124	4.989	6.328	8.401	12.480	14.147

From Table 6.6 we observe that h increases as n increases and α decreases. As pointed out by Sullivan and Woodall (1996), it is not important for the preliminary application to find exact control limits that correspond to a specific *FAP*. Instead it is sufficient to use computationally convenient limits having approximately the desired performance. Consequently, ZZW derived a formula to approximate the decision interval h:

$$h = (-0.0905n + 0.5221)\log\alpha - 0.1923n + 1.0248. \qquad (6.25)$$

Using equation (6.25) to approximate the decision interval we obtain the following values for h.

Table 6.7. Approximated h values for the CUSUM Mann-Whitney chart[†].

	α							
n	**0.07**	**0.06**	**0.05**	**0.04**	**0.03**	**0.02**	**0.01**	**0.0075**
20	0.604	0.802	1.037	1.324	1.695	2.217	3.110	3.480
30	1.087	1.425	1.825	2.314	2.945	3.834	5.354	5.985
40	1.571	2.048	2.613	3.305	4.196	5.452	7.599	8.490
50	2.055	2.672	3.401	4.295	5.446	7.069	9.844	10.995
60	2.538	3.295	4.190	5.285	6.697	8.687	12.089	13.500

Comparing the approximated h values with the simulated results in Table 6.6 it is clear that the approximated decision interval using equation (6.25) performs very well as they agree well with the values of Table 6.6.

6.2.3. Performance comparison

The performance of a control chart is usually judged in terms of certain characteristics associated with its run-length distribution. In a Phase I setting, the *FAP*, which is the probability of at least one false alarm out of many comparisons, is used for performance comparison as opposed to using the *FAR*, which is the probability of a single false alarm

[*] Table 6.6 appears in ZZW, page 5, Table 1.
[†] The values in Table 6.7 were generated using Microsoft Excel.

involving only a single comparison. By using the *FAP* we take into account that the signaling events are dependent. ZZW looked at the *FAP*, the true signal probability (*TSP*) and the average true signal probability (*ATSP*) for performance comparison*. The *TSP* is the probability of a signal when a shift has occurred. The *ATSP* is defined as $ATSP = \sum_{k=1}^{n-1} F(k)TSP_k$ where TSP_k denotes the *TSP* of a control chart when the shift occurs after the k^{th} observation and $F(k)$ denotes the distribution of the position of the shifts. To ensure fair comparison between two charts, charts with the same *FAP* are considered and the chart with the larger *TSP* (or *ATSP*) is the preferred chart. In their paper, ZZW assumes that the position of the shift is uniformly distributed so that the position of the shift is equally likely at any point and, under this assumption, the CUSUM Mann-Whitney chart, the CUSUM chart for detecting the linear trend (CUSUM LT) chart (see Koning and Does (2000)) and the likelihood ratio test (LRT) chart (see Sullivan and Woodall (1996)) are compared. For the LRT and CUSUM LT charts the assumption of normality is necessary, whereas with the CUSUM Mann-Whitney chart no assumption about the underlying process distribution needs to be made. The performances of these charts are compared for five distributions, namely the Normal, Chi-square, Student *t*, Weibull and Lognormal, respectively. We would expect to find that the CUSUM Mann-Whitney chart performs better compared to the CUSUM LT and LRT charts when the distribution is skewed or heavy-tailed.

(i) The standard normal distribution

For the standard normal distribution both a step shift and a linear trend shift are used to evaluate chart performance. Recall that charts with the same *FAP* ($= 0.05$) are considered to ensure fair comparison and that the chart with the larger *ATSP* is the preferred chart.

* The terms *TSP* and *ATSP* are fairly new and are introduced by Sullivan and Woodall (1996).

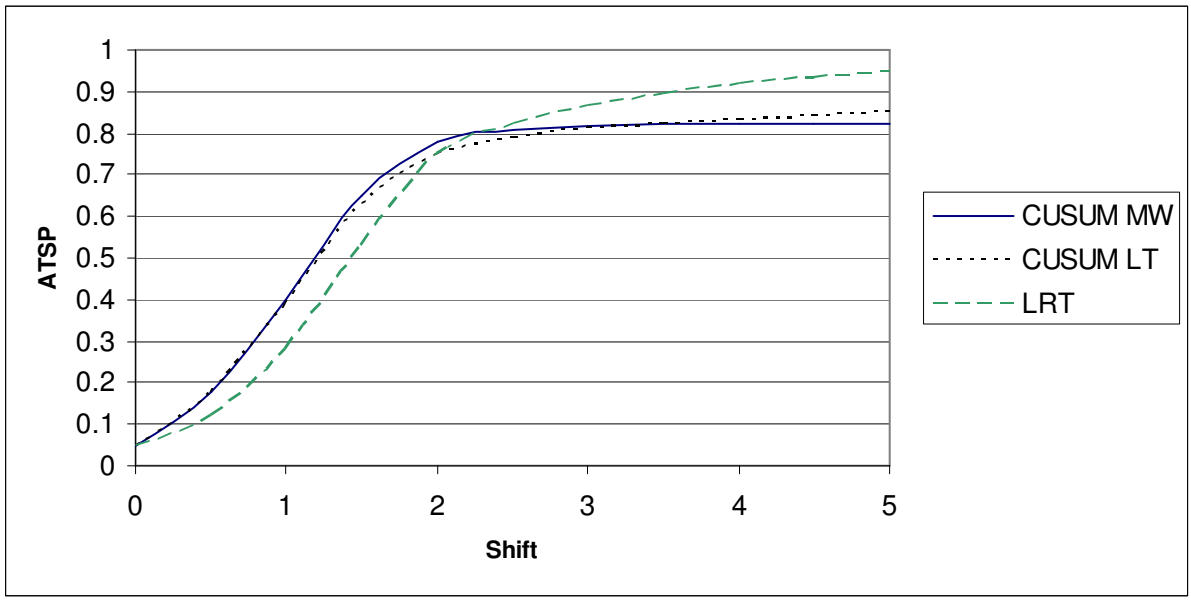

Figure 6.6. The *ATSP* values for a single step shift when the data is from a $N(0,1)$ distribution.

When comparing the CUSUM Mann-Whitney chart with the CUSUM LT chart we find that these charts have comparable performance. When comparing all three charts we find that the CUSUM Mann-Whitney chart has a slight disadvantage in detecting large shifts.

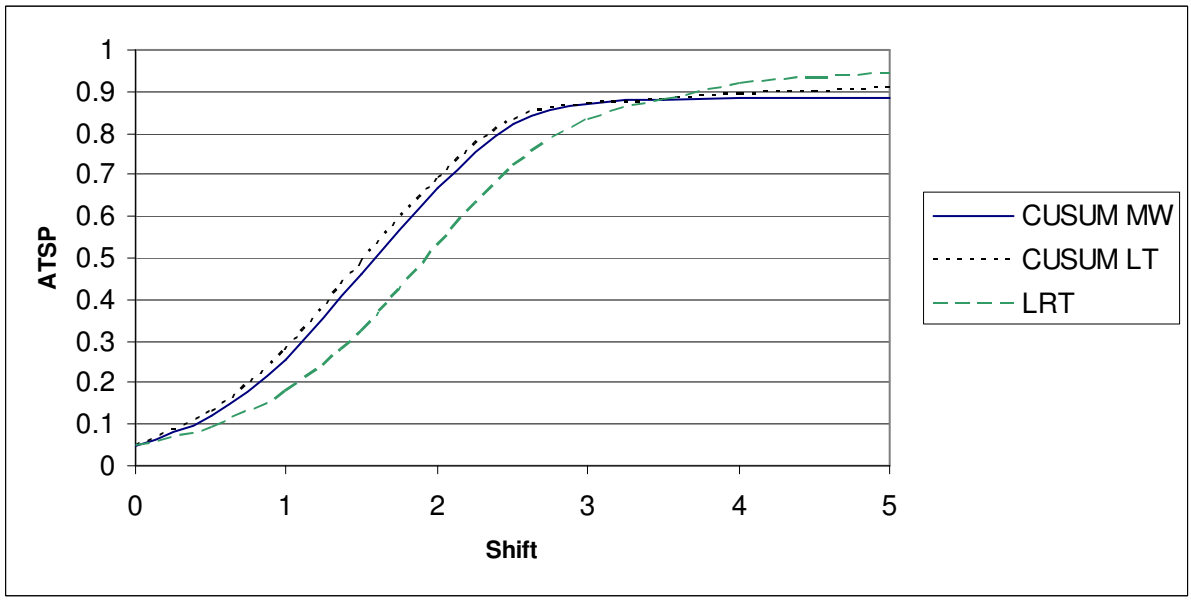

Figure 6.7. The *ATSP* values for a linear trend shift when the data is from a $N(0,1)$ distribution.

When comparing the CUSUM Mann-Whitney chart with the CUSUM LT chart we find that the CUSUM LT chart is performing slightly better than the CUSUM Mann-Whitney chart. When comparing all three charts we find that the CUSUM Mann-Whitney chart has a slight disadvantage in detecting large shifts. It is worth mentioning that even for normally distributed data the CUSUM Mann-Whitney chart is performing very well. The performance of the CUSUM Mann-Whitney chart could be improved by changing the reference value to some other value (recall that ZZW set the reference value equal to 2).

(ii) The *t*-distribution

The *t*-distribution with degrees of freedom 2 is symmetric around zero and as the number of degrees of freedom increases, the difference between the *t*-distribution and the standard normal distribution becomes smaller and smaller.

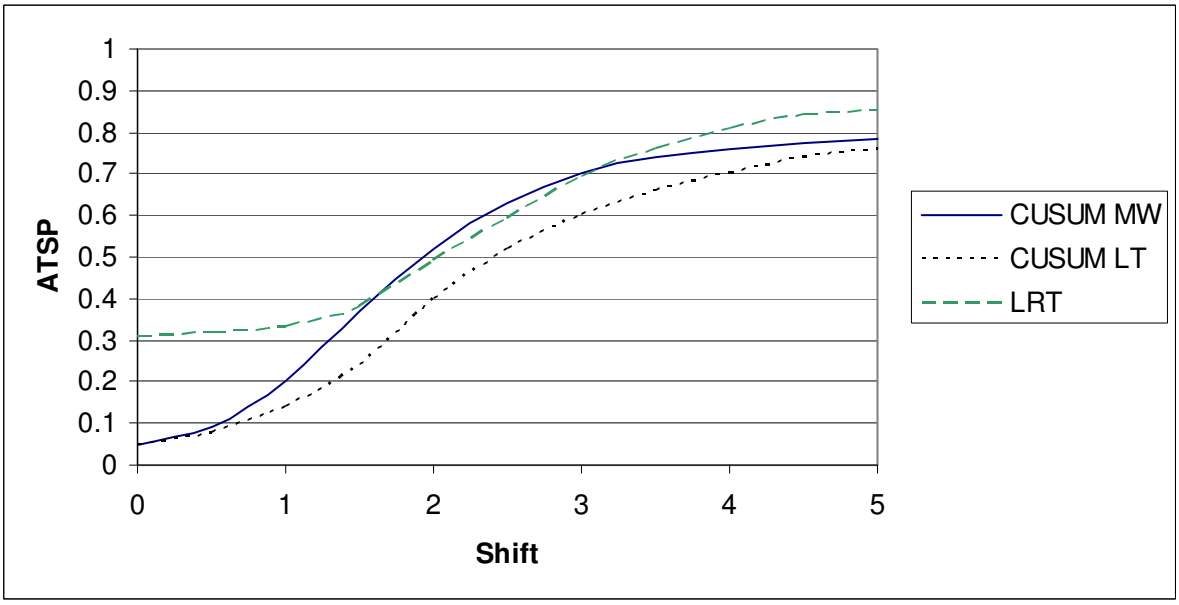

Figure 6.8. The *ATSP* values for a single step shift when the data is from a $t(2)$ distribution.

Recall that charts with the same *FAP* $(=0.05)$ are considered to ensure fair comparison and that the chart with the larger *ATSP* is the preferred chart. From Figure 6.8 we can see that the LRT chart can not obtain the specified *FAP* of 0.05. Consequently, the LRT chart is not compared to the other charts under $t(2)$ shift alternatives. When comparing the CUSUM Mann-Whitney chart with the CUSUM LT chart we find that the CUSUM Mann-Whitney chart is performing better than the CUSUM LT chart, since the *ATSP* values for the

CUSUM Mann-Whitney chart are larger than that of the CUSUM LT chart. It should be noted that the differences are relatively small over all values of the shift.

(iii) The Chi-square distribution

The Chi-square distribution is highly skewed to the right and as a result we would expect the performance of the CUSUM Mann-Whitney chart to be better than that of the CUSUM LT and LRT charts (since they have the additional assumption of normality).

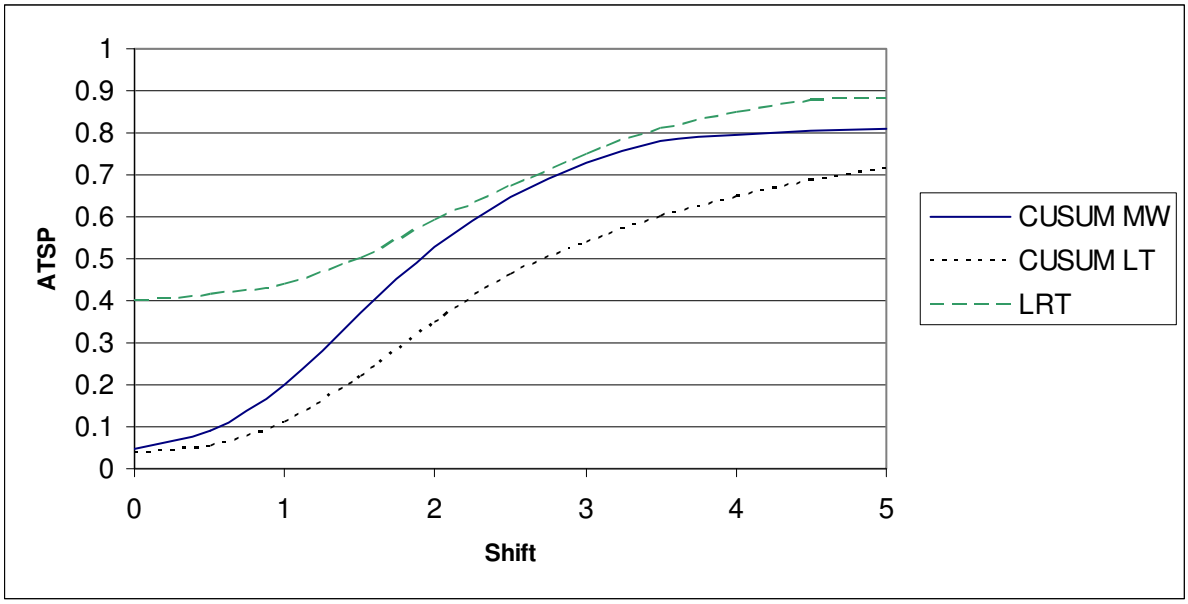

Figure 6.9. The *ATSP* values for a single step shift when the data is from a $\chi^2(2)$ distribution.

Similar to the previous comparison, we see that the LRT chart can not obtain the specified *FAP* of 0.05. Consequently, the LRT chart is not compared to the other charts under $\chi^2(2)$ shift alternatives. When comparing the CUSUM Mann-Whitney chart with the CUSUM LT chart we find that the CUSUM Mann-Whitney chart is performing better than the CUSUM LT chart, since the *ATSP* values for the CUSUM Mann-Whitney chart are larger than that of the CUSUM LT chart. It should be noted that the differences are larger than those under $t(2)$ shift alternatives.

(iv) The Weibull distribution

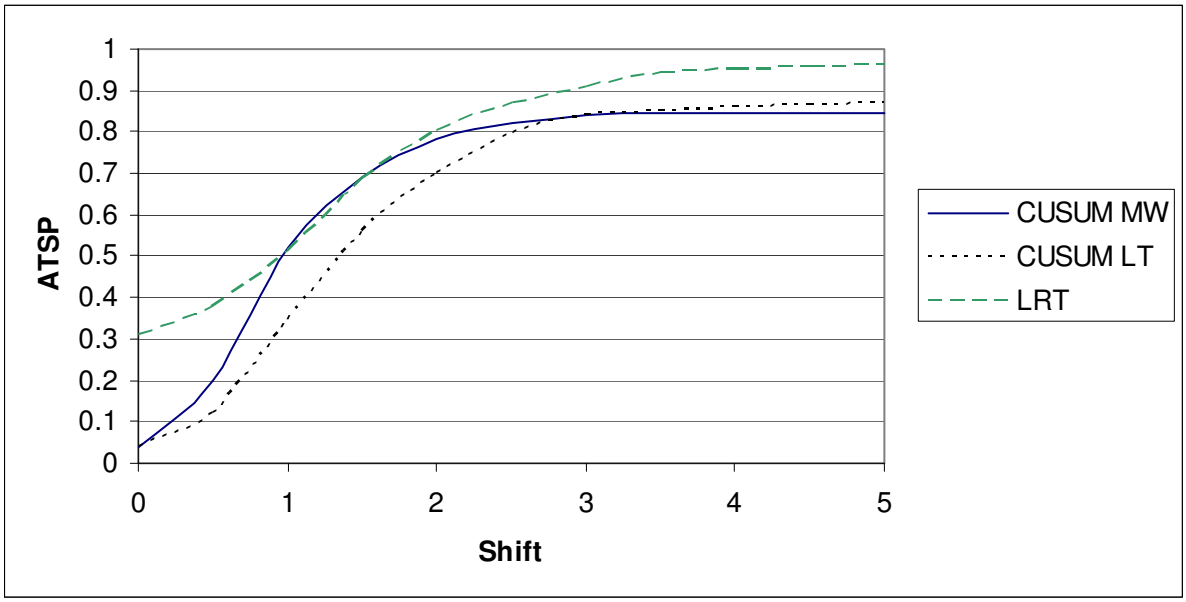

Figure 6.10. The *ATSP* values for a single step shift when the data is from a Weibull(1,1) distribution.

Similar to the previous two comparisons, we see that the LRT chart can not obtain the specified *FAP* of 0.05. Consequently, the LRT chart is not compared to the other charts under Weibull(1,1) shift alternatives. When comparing the CUSUM Mann-Whitney chart with the CUSUM LT chart we find that the CUSUM Mann-Whitney chart is performing better than the CUSUM LT chart for small shift sizes, whereas, for large shift sizes the opposite is true, although to a very small extent.

(v) The lognormal distribution

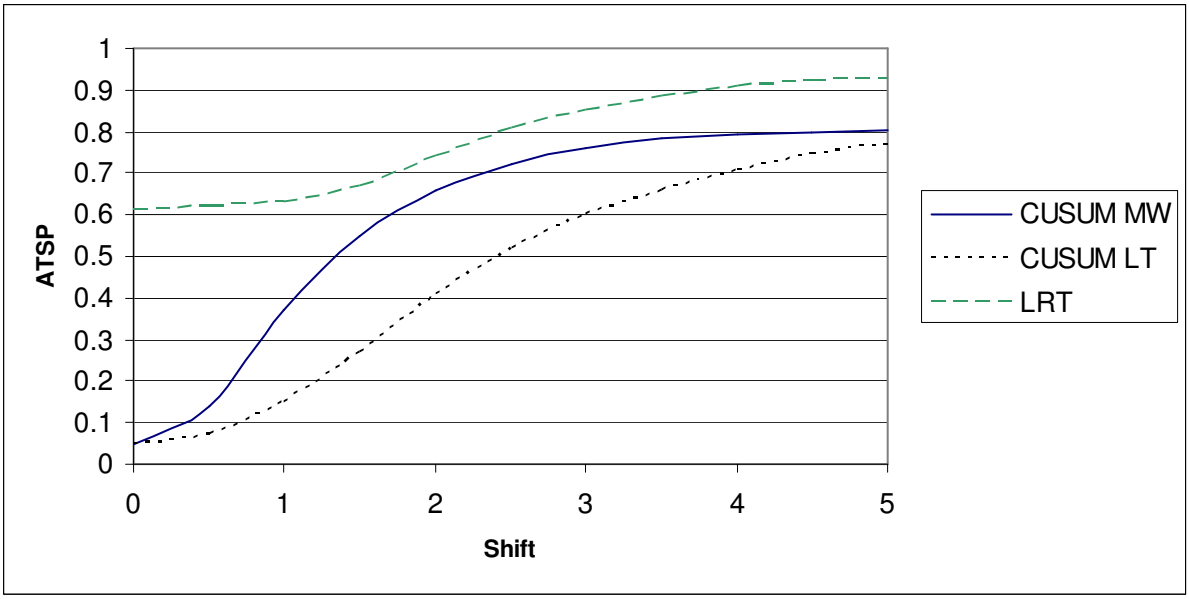

Figure 6.11. The *ATSP* values for a single step shift when the data is from a lognormal(0,1) distribution.

Similar to the previous three comparisons, we see that the LRT chart can not obtain the specified *FAP* of 0.05. Consequently, the LRT chart is not compared to the other charts under lognormal(0,1) shift alternatives. When comparing the CUSUM Mann-Whitney chart with the CUSUM LT chart we find that the CUSUM Mann-Whitney chart is performing better than the CUSUM LT chart, since the *ATSP* values for the CUSUM Mann-Whitney chart are larger than that of the CUSUM LT chart.

Table 6.8. A summary of the performances of the CUSUM Mann-Whitney, CUSUM LT and LRT charts for five different distributions.

Distribution	**Type of shift**	**Preferred control chart***
Normal(0,1)	Linear trend shift	For shifts < 3.5: 1) CUSUM LT 2) CUSUM MW 3) LRT For shifts > 3.5: 1) LRT 2) CUSUM LT 3) CUSUM MW
Normal(0,1)	Single step shift	For shifts < 2.5: 1) CUSUM MW or CUSUM LT (comparable performance) 2) LRT For shifts > 2.5: 1) LRT 2) CUSUM LT or CUSUM MW (comparable performance)
$t(2)$	Single step shift	1) CUSUM MW 2) CUSUM LT
$\chi^2(2)$	Single step shift	1) CUSUM MW 2) CUSUM LT
Weibull(1,1)	Single step shift	For shifts < 3.0: 1) CUSUM MW 2) CUSUM LT For shifts > 3.0: 1) CUSUM LT 2) CUSUM MW
Lognormal(0,1)	Single step shift	1) CUSUM MW 2) CUSUM LT

Example 6.2

A CUSUM Mann-Whitney control chart

We illustrate the CUSUM Mann-Whitney control chart using a set of simulated data used by Sullivan and Woodall (1996; Table 2) and ZZW (2007; Table 2). This data set is ideal for use in this Phase I problem, since it is known to have a single step shift in the mean. There are 30 observations, i.e. $n = 30$, which are distributed as follows: $X_i \sim N(0,1)$ for $i \leq 15$ and $X_i \sim N(1,1)$ for $i > 15$ (a value of 1 was added to the last 15 observations causing the data

* The control charts are ranked from the most preferred to the least preferred. The LRT chart could not be compared to the CUSUM MW and CUSUM LT charts under $t(2)$, $\chi^2(2)$, Weibull(1,1) and Lognormal(0,1) shift alternatives, since the LRT chart could not obtain the specified *FAP* of 0.05.

to exhibit a step shift in the middle of the sample). Clearly, there is a known change-point at $t = 15$ where the mean has shifted from 0 to 1. The Mann-Whitney statistics (MW_t), the corresponding expected values ($E(MW_t)$), standard deviations ($stdev(MW_t)$) and standardized values (SMW_t), respectively, are given in Table 6.9. The CUSUM S_i^+ and S_i^- values are also given in Table 6.9 and illustrated in Figure 6.12. The starting values are set equal to zero, that is, $S_0^+ = S_0^- = 0$ (as recommended by Page (1954)).

Table 6.9. Data and calculations for the CUSUM Mann-Whitney chart when $k = 2$.*

i	X_i	MW_t	$E(MW_t)$	$stdev(MW_t)$	SMW_t	S_i^+	S_i^-
1	-0.69	6	14.5	8.655	-0.982	0.000	0.000
2	0.56	20	28.0	12.028	-0.665	0.000	0.000
3	-0.96	23	40.5	14.465	-1.210	0.000	0.000
4	-0.11	29	52.0	16.391	-1.403	0.000	0.000
5	-0.25	33	62.5	17.970	-1.642	0.000	0.000
6	0.45	41	72.0	19.287	-1.607	0.000	0.000
7	-0.26	42	80.5	20.394	-1.888	0.000	0.000
8	0.68	54	88.0	21.323	-1.595	0.000	0.000
9	0.22	57	94.5	22.096	-1.697	0.000	0.000
10	-2.10	49	100.0	22.730	-2.244	0.000	-0.244
11	0.65	56	104.5	23.236	-2.087	0.000	-0.331
12	-1.49	47	108.0	23.622	-2.582	0.000	-0.913
13	-2.49	35	110.5	23.894	-3.160	0.000	-2.073
14	-1.11	25	112.0	24.055	-3.617	0.000	-3.690
15	0.23	23	112.5	24.109	-3.712	0.000	-5.402
16	2.16	35	112.0	24.055	-3.201	0.000	-6.603
17	1.95	45	110.5	23.894	-2.741	0.000	-7.344
18	1.54	52	108.0	23.622	-2.371	0.000	-7.715
19	0.67	52	104.5	23.236	-2.259	0.000	-7.974
20	1.09	54	100.0	22.730	-2.024	0.000	-7.998
21	1.37	56	94.5	22.096	-1.742	0.000	-7.740
22	0.69	55	88.0	21.323	-1.548	0.000	-7.288
23	2.26	61	80.5	20.394	-0.956	0.000	-6.244
24	1.86	63	72.0	19.287	-0.467	0.000	-4.711
25	0.62	55	62.5	17.970	-0.417	0.000	-3.128
26	-1.04	34	52.0	16.391	-1.098	0.000	-2.226
27	2.30	37	40.5	14.465	-0.242	0.000	-0.468
28	0.07	20	28.0	12.028	-0.665	0.000	0.000
29	1.49	15	14.5	8.655	0.058	0.000	0.000
30	0.52						

* See SAS Program 9 in Appendix B for the calculation of the values in Table 6.9. This table also appears in ZZW, page 7, Table 2.

Table 6.9 also appears in ZZW, page 7, Table 2. It should be noted that the CUSUM S_i^- values that we obtained in Table 6.9 are different from those in ZZW, since they used equation (6.24) to calculate S_i^-, whereas we used equation (6.23).

As illustration, the expected value ($E(MW_t)$), standard deviation ($stdev(MW_t)$), standardized value (SMW_t), CUSUM S_i^+ and S_i^- values will be calculated for $t = 1$.

$$E(MW_1) = \frac{1(30-1)}{2} = 14.5, \quad stdev(MW_t) == \sqrt{\frac{1(30-1)(30+1)}{12}} = 8.655,$$

$$SMW_t = \frac{MW_t - E(MW_t)}{stdev(MW_t)} = \frac{6-14.5}{8.655} = -0.982,$$

$$S_1^+ = \max[0, S_0^+ + SMW_1 - k] = \max[0, 0 + (-0.982) - 2] = 0,$$

$$S_1^- = \min[0, S_0^- + SMW_1 + k] = \min[0, 0 + (-0.982) + 2] = 0.$$

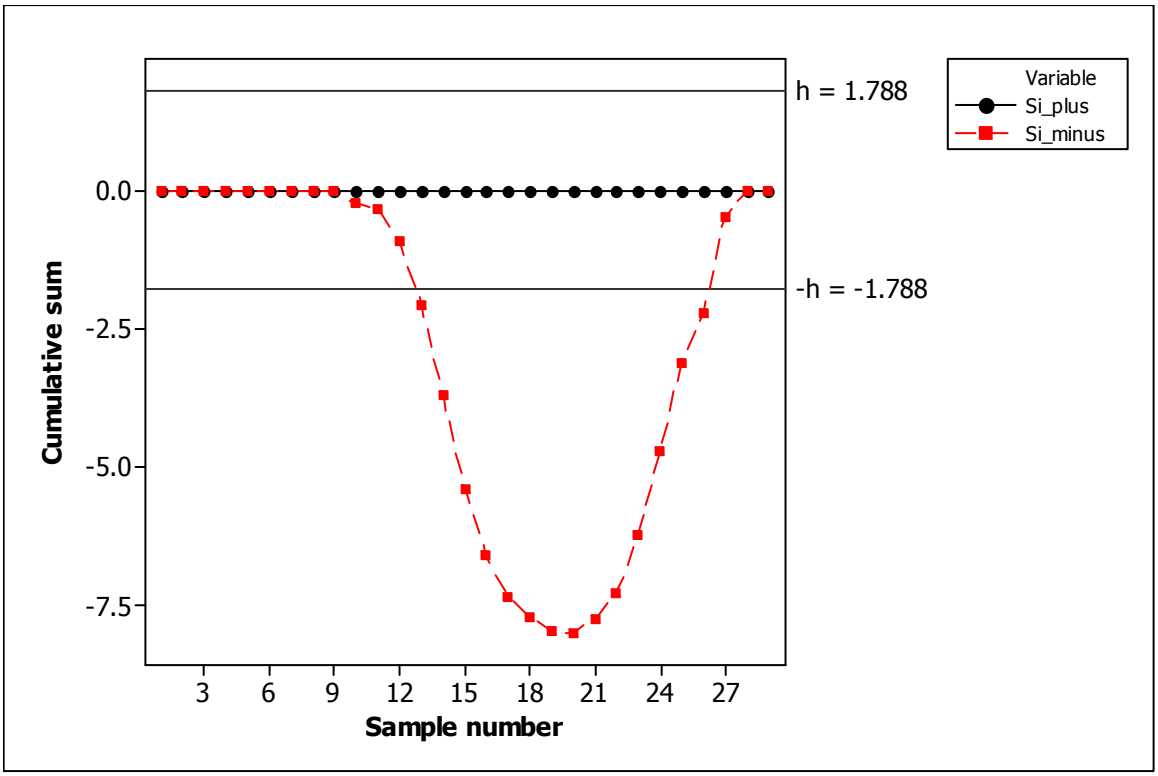

Figure 6.12. The CUSUM Mann-Whitney chart with $n = 30$, $k = 2$ and $h = 1.788$.

For a sample size of 30 and a desired *FAP* of 0.05, the decision interval is taken to be 1.788 (see Table 6.6). From Figure 6.12 we see that the process is out-of-control starting at sample number 13 using the CUSUM Mann-Whitney chart, whereas the LRT chart of

Sullivan and Woodall (1996) indicated that observation 15 is the most likely location of the shift. Hence, the CUSUM Mann-Whitney chart detects that the mean has shifted upwards.

6.2.4. Summary

ZZW found that the Phase I CUSUM Mann-Whitney chart has good performance compared to the CUSUM LT chart for all distributions, except for the Weibull distribution for large shifts. Their proposed nonparametric chart for preliminary analysis can be useful for quality practitioners in applications where not much is known or can be assumed about the process distribution. Although a lot has been accomplished in the last few years regarding the development of control charts based on the Mann-Whitney statistic, more remains to be done. In terms of research, work needs to be done on a Phase II CUSUM-type chart based on the Mann-Whitney statistic for individual observations and subgroups (recall that the control chart proposed by ZZW is a Phase I CUSUM-type chart for preliminary analysis of individual observations). Also recall that ZZW assumes that the position of the shift is uniformly distributed. One could, for future research, consider other distributions for the position of the shift. Furthermore, work needs to be done on Phase I and Phase II EWMA-type charts based on the Mann-Whitney statistic. Clearly, there are lots of opportunities for future research.

Chapter 7: Concluding remarks

In this thesis, we mentioned some of the key contributions and ideas and a few of the more recent developments in the area of univariate nonparametric control charts. We considered the three main classes of control charts: the Shewhart, CUSUM and EWMA control charts and their refinements. The statistics used in nonparametric control charts are mostly signs, ranks and signed-ranks and related to nonparametric procedures, such as the Wilcoxon signed-rank test and the Mann-Whitney-Wilcoxon rank-sum test. We described the sign and signed-rank control charts under each of the three classes in Chapters 2 and 3, respectively. In Chapter 4 we only considered the Shewhart-type sign-like control chart, since the CUSUM- and EWMA-type control charts have not been developed for the sign-like case. In Chapter 5 we only considered the Shewhart-type signed-rank-like control chart, since the CUSUM- and EWMA-type control charts have not been developed for the signed-rank-like case. Finally, in Chapter 6 we only considered the Shewhart- and CUSUM-type Mann-Whitney-Wilcoxon control charts, since the EWMA-type control chart has not been developed for the Mann-Whitney-Wilcoxon statistic. Clearly, there are lots of opportunities for future research.

We considered decision problems under both Phase I and Phase II (see Section 1.5 for a distinction between the two phases). In all the sections of this thesis we considered Phase II process monitoring, except in Section 6.2 where a CUSUM-type control chart for the preliminary Phase I analysis of individual observations based on the Mann-Whitney two-sample test is proposed. Although the field of preliminary Phase I analysis is interesting and the body of literature on Phase I control charts is growing, more research is necessary on Phase I nonparametric control charts in general.

We only discussed univariate nonparametric control charts designed to track the location of a continuous process, since very few charts are available for scale. Therefore, future research needs to be done on monitoring the scale and simultaneously monitoring the location and the scale of a process.

There has been other work on nonparametric control charts. Among these, for example, Albers and Kallenberg (2004) studied conditions under which the nonparametric

charts become viable alternatives to their parametric counterparts. They consider Phase II charts for individual observations in case U based on empirical quantiles or order statistics. The basic problem is that for the very small *FAR* typically used in the industry, a very large reference sample size is usually necessary to set up the chart. They discuss various remedies for this problem.

Another area that has received some attention is control charts for variable sampling intervals (VSI). In a typical control charting environment, the time interval between two successive samples is fixed, and this called a fixed sampling interval (FSI) scheme. VSI schemes allow the user to vary the sampling interval between taking samples. This idea has intuitive appeal since when one or more charting statistics fall close to one of the control limits but not quite outside, it seems reasonable to sample more frequently, whereas when charting statistics plot closer to the centerline, no action is necessary and only a few samples might be sufficient. On the point of VSI control schemes see for example, Amin (1987), Reynolds et al (1990), Rendtel (1990), Saccucci et al (1992) and Amin and Hemasinha (1993). These researchers examined combining the VSI approach with the Shewhart, CUSUM and the EWMA control schemes, respectively. They demonstrated that the VSI control schemes are more efficient than the corresponding FSI control schemes. VSI control schemes use a long sampling interval between successive samples when the plotting statistic is close to target and a shorter sampling interval otherwise. Initially, the short sampling interval could be used for the first few samples to offer protection at start-up. Amin and Widmaier (1999) compared the Shewhart $\overline{X}$ charts with sign control charts, under the FSI and VSI schemes, on the basis of *ARL* for various shift sizes and several underlying distributions like the normal distribution and distributions that are heavy-tailed and/or asymmetric like the double exponential and the gamma. It is seen that the nonparametric VSI sign charts are more efficient than the corresponding FSI sign charts.

We hope this thesis leads to a wider acceptance of nonparametric control charts among practitioners and promotes further interest in the development of nonparametric control charts.

Appendix A: Theorems and proofs

Theorem 1:

The *ARL* of the two-sided chart can be expressed as a function of the average run lengths of the one-sided charts through the expression

$$\frac{1}{ARL} = \frac{1}{ARL^U} + \frac{1}{ARL^L} \tag{A1}$$

where ARL^U and ARL^L denotes the average run lengths for the upper and lower one-sided charts, respectively. This result applies to both Shewhart- and CUSUM-type charts. A proof of expression (A1) is given by using the properties of generating functions.

Proof to Theorem 1:

Generating functions

Let X be a random variable whose possible values are restricted to the nonnegative integers $\{0,1,2,...\}$ and write $c_j = P(X = j)$. The probability generating function (hereafter pgf) is defined as

$$\Pi_X(s) = \sum_{j=0}^{\infty} c_j s^j = \sum_{j=0}^{\infty} P(X = j)s^j = P(X = 0)s^0 + P(X = 1)s^1 + P(X = 2)s^2 + ...$$

where s must be restricted to a region in which the power series is convergent. The power series always converges if

$$|s| \le 1, \text{ that is, } -1 \le s \le 1. \tag{A2}$$

An alternative definition of $\Pi_X(s)$ is

$$\Pi_X(s) = E\left(s^X\right). \tag{A3}$$

Properties of generating functions

Let

$$q_j = c_{j+1} + c_{j+2} + ... = P(X = j+1) + P(X = j+2) + ... = P(X > j) \text{ for } j = 0,1,2,... \tag{A4}$$

be the 'tail' probabilities. Then

$$\sum_{j=0}^{\infty} q_j = (c_1 + c_2 + ...) + (c_2 + c_3 + ...) + ... = c_1 + 2c_2 + 3c_3 + ... = \sum_{j=1}^{\infty} jc_j = \sum_{j=0}^{\infty} jc_j = E(X).$$

Therefore the sequence $\{q_j\}$ for $j = 0,1,2,...$ is of importance, because it constitutes another probability distribution on the integers $0,1,2,...$ with its pgf given by

$$\sum_{j=0}^{\infty} q_j s^j = \frac{1 - \sum_{j=0}^{\infty} c_j s^j}{1 - s} \text{ for } -1 < s < 1 \qquad \text{(A5)}$$

which simplifies to

$$Q(s) = \frac{1 - \Pi_X(s)}{1 - s} \text{ for } -1 < s < 1. \qquad \text{(A6)}$$

The condition $-1 < s < 1$ is due to the convergence rule (A2), and the fact that expressions (A5) and (A6) will only hold for $s \neq 1$.

Signalling event

Let ε denote a signalling event. Then ε is a *recurrent* event, because if ε occurs on the j^{th} trial, we treat trial $j+1$ as though it were the first trial. Let the random variable Y denote the number of the trial on which event ε occurs for the first time. Let q_j denote the probability of no occurrences in the first j trials of an event ε. Then $q_j = P(\varepsilon \text{ does not occur in the first } j \text{ trials}) = P(Y > j)$. If $Y > j$, there have been no indications of a changed process in the first j points. Let c_j denote the probability that an event ε occurs for the *first time* on the j^{th} trial. Let p_j denote the probability that an event ε occurs on the j^{th} trial. The set of initial conditions is:

$q_0 = P(\varepsilon \text{ does not occur in the first 0 trials}) = 1$	(A7)
$c_0 = 0$	(A8)
$p_0 = 1$	(A9)

Let $Q(s), C(s)$ and $P(s)$ denote the generating functions for the probabilities q_j, c_j and p_j, respectively. The pgf uniquely determines the corresponding probability distribution

and in turn the probability distribution on $0,1,2,\ldots$, uniquely determines the pgf. Clearly, there is a 1-1 correspondence between the probability distributions and the pgf's.

Generating function $Q(s)$:

$$Q(s) = \sum_{j=0}^{\infty} q_j s^j = q_0 s^0 + q_1 s^1 + q_2 s^2 + q_3 s^3 + \ldots$$

Applying initial condition (A7), we have that

$$Q(s) = 1 + q_1 s + q_2 s^2 + q_3 s^3 + \ldots$$

The generating function is helpful in obtaining moments of the distribution of Y. Particularly,

$$Q(1) = 1 + q_1 + q_2 + q_3 + \ldots = 1 + P(Y > 1) + P(Y > 2) + P(Y > 3) + \ldots = E(Y) \qquad \text{(A10)}$$

where $E(Y)$ denotes the average number of trials between consecutive occurrences of signalling events.

Generating function $C(s)$:

$$C(s) = \sum_{j=0}^{\infty} c_j s^j = c_0 s^0 + c_1 s^1 + c_2 s^2 + c_3 s^3 + \ldots$$

Applying initial condition (A8), we have that

$$C(s) = c_1 s + c_2 s^2 + c_3 s^3 + \ldots$$

Due to the fact that we only consider recurrent events which have finite recurrence times, we have that

$$C(1) = c_1 + c_2 + c_3 + \ldots = 1. \qquad \text{(A11)}$$

Generating function $P(s)$:

$$P(s) = \sum_{j=0}^{\infty} p_j s^j = p_0 s^0 + p_1 s^1 + p_2 s^2 + p_3 s^3 + \ldots$$

Applying initial condition (A9), we have that

$$P(s) = 1 + p_1 s + p_2 s^2 + p_3 s^3 + \ldots$$

Corollary A1:

The $p's$ can be determined in terms of the $c's$:

$$P(s) = \frac{1}{1 - C(s)}$$

Proof:

Note: In this proof the set of initial conditions holds (see (A7), (A8) and (A9)).

The equation given below holds, since ε is a recurrent event:

$$p_j = c_j p_0 + c_{j-1} p_1 + c_{j-2} p_2 + ... + c_1 p_{j-1} + c_0 p_j$$

For $j = 0$: $p_0 = c_0 p_0$

For $j = 1$: $p_1 = c_1 p_0 + c_0 p_1$

For $j = 2$: $p_2 = c_2 p_0 + c_1 p_1 + c_0 p_2$

Continuing this way for $j = 3,4,...$ we have that

$$p_0 + p_1 + p_2 + ... = (c_0 p_0) + (c_1 p_0 + c_0 p_1) + (c_2 p_0 + c_1 p_1 + c_0 p_2) + ...$$

$$= (p_0 + p_1 + p_2 + ...)(c_0 + c_1 + c_2 + ...)$$

Applying initial conditions (A8) and (A9) we have that

$$1 + p_1 + p_2 + ... = (1 + p_1 + p_2 + ...)(0 + c_1 + c_2 + ...) \quad \text{(A12)}$$

Multiplying (A12) through by s^j and summing over j from one to infinity, we obtain

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} p_k \sum_{i=0}^{\infty} c_i$$

$$\sum_{j=1}^{\infty} \sum_{k=0}^{\infty} p_k s^j = \sum_{j=1}^{\infty} s^j \sum_{k=0}^{\infty} p_k \sum_{i=0}^{\infty} c_i$$

$$\sum_{j=0}^{\infty} p_j s^j = \sum_{j=0}^{\infty} p_j s^j \sum_{k=0}^{\infty} c_k s^k + 1$$

$$P(s) = P(s)C(s) + 1 \quad \text{(A13)}$$

(refer to expression (A11) for an explanation of why the constant appears in expression (A13)) . Re-writing expression (A13) we obtain

$$P(s) = \frac{1}{1 - C(s)} \quad \text{(A14)}$$

Corollary A2:

The q's can be determined in terms of the c's:

$$Q(s) = \frac{1 - C(s)}{1 - s} \quad \text{for} \quad -1 < s < 1 .$$

Proof:

By using equations (A5) and (A6), the q's can be determined in terms of the c's:

$$\sum_{j=0}^{\infty} q_j s^j = \frac{1 - \sum_{j=0}^{\infty} c_j s^j}{1 - s} \tag{A15}$$

which simplifies to

$$Q(s) = \frac{1 - C(s)}{1 - s} \quad \text{for} \quad -1 < s < 1 . \tag{A16}$$

From corollary A1 and corollary A2 we have

$$P(s) = \frac{1}{1 - C(s)}$$

and

$$\frac{1}{1 - P(s)} = \frac{1}{(1 - s)Q(s)} \quad \text{for} \quad -1 < s < 1$$

Therefore, we have that

$$P(s) = \frac{1}{1 - C(s)} = \frac{1}{(1 - s)Q(s)} \quad \text{for} \quad -1 < s < 1 \tag{A17}$$

which can be re-written as

$$(1 - s)P(s) = \frac{1}{Q(s)} \quad \text{for} \quad -1 < s < 1 \tag{A18}$$

Taking the limit of $(1 - s)P(s)$ as s approaches unity and applying (A10) we have that

$$\lim_{s \to 1}(1 - s)P(s) = \lim_{s \to 1} \frac{1}{Q(s)} = \frac{1}{Q(1)} = \frac{1}{E(Y)} . \tag{A19}$$

Let ε_U and ε_L denote the signalling events for the upper and lower one-sided charts, respectively. Let ε_{UL} denote a signalling event of the two-sided chart. Let $E_L(Y)$, $E_U(Y)$ and $E_{UL}(Y)$ denote the average recurrence time of ε_L, ε_U and ε_{UL}, respectively. We would like

to determine $E_{UL}(Y)$ from $E_L(Y)$ and $E_U(Y)$. Either ε_L or ε_U (but not both) occurs on every trial on which ε_{UL} occurs, and this leads to equation (A20)

$$P_{UL,j} = P_{L,j} + P_{U,j} \tag{A20}$$

Multiplying (A20) through by s^j and summing over j from one to infinity, we obtain the generating functions:

$$P_{UL}(s) = P_L(s) + P_U(s) - 1 \tag{A21}$$

The constant appears, because probabilities sum to unity. Summing over j from one to infinity over the probabilities on the left-hand side of equation (A20), must equal one. Summing over j from one to infinity over the probabilities of the first term on the right-hand side of equation (A20), must equal one *and* summing over j from one to infinity over the probabilities of the second term on the right-hand side of equation (A20), must equal one. Therefore, summing the two terms on the right-hand side equals two. The problem arises: The left-hand side of the equation equals one while the right-hand side of the same equation equals two. This problem is solved by subtracting one from the right-hand side of the equation.

Multiplying (A21) through by $(1-s)$ and taking the limit as s approaches unity, we have that $\lim_{s\to 1}(1-s)P_{UL}(s) = \lim_{s\to 1}(1-s)P_L(s) + \lim_{s\to 1}(1-s)P_U(s) - \lim_{s\to 1}(1-s)$. From (A19) we have that

$$\frac{1}{E_{UL}(Y)} = \frac{1}{E_L(Y)} + \frac{1}{E_U(Y)} - (1-1)$$

which simplifies to

$$\frac{1}{E_{UL}(Y)} = \frac{1}{E_L(Y)} + \frac{1}{E_U(Y)}. \tag{A22}$$

Theorem 2:

Fu, Spiring and Xie (2002) defined the moment generating function (hereafter mgf) as $\phi(t) = 1 + (e^t - 1)\underline{\xi}(I - e^t Q)^{-1}\underline{1}$ and used the mgf to obtain expressions for the first and second moments of the run length variable N. Fu and Lou (2003) defined the probability generating function (hereafter pgf) as $\varphi(t) = 1 + (t-1)\underline{\xi}(I - tQ)^{-1}\underline{1}$ and used the pgf to obtain expressions for the first and second moments of the run length variable N. Although they used different methods, both were able to obtain the following expressions for the first and second moments of the run length variable N:

$$E(N) = \underline{\xi}(I - Q)^{-1}\underline{1}$$

$$E(N^2) = \underline{\xi}(I + Q)(I - Q)^{-2}\underline{1}$$

where Q is the matrix that contains all the transition probabilities of going from a non-absorbing state to a non-absorbing state, I is the identity matrix, $\underline{\xi} = (1,0,0,...,0)$ is a row vector with 1 at the 1^{st} element and zeros elsewhere, $\underline{1} = (1 \quad ... \quad 1)$ is a column vector with all elements equal to unity (refer to the Section 2.3 for more detail about the construction and the dimensions of these matrices).

In this appendix the derivation of the first and second moments of the run length variable N will be done using both the mgf and the pgf.

Proof to Theorem 2:

A power series is defied as $f(x) = \sum_{n=0}^{\infty} a_n x^n$. It is also referred to as the generating function. Generating functions are very useful combinatorial enumeration problems. In general we have that $x(1-x)^{-1} = x(1 + x + x^2 + ...)$ for $|x| < 1$. Similarly, in this example we will use the fact that $Q(I - Q)^{-1} = Q(I + Q + Q^2 + ...)$. In addition, generally we have that $x(1-x)^{-2} = x + 2x^2 + 3x^3 + ... = x(1 + 2x + 3x^2 + ...)$ for $|x| < 1$. Similarly, in this example we will use the fact that $Q(I - Q)^{-2} = Q(I + 2Q + 3Q^2 + ...)$.

Moment generating function

(See page 373 of Fu, Spiring and Xie (2002))

The mgf is given by $\phi(t) = 1 + (e^t - 1)\underline{\xi}(I - e^t Q)^{-1}\underline{1}$. It is well-known that if the mgf is differentiable at zero, then the n^{th} moment is given by $\phi^{(n)}(0)$. Therefore, in order to find the first moment, we will have to calculate the first order derivative of the mgf in the point $t = 0$, that is, $E(N) = \phi'(0)$. Similarly, in order to find the second moment, we will have to calculate the second order derivative of the mgf in the point $t = 0$, that is, $E(N^2) = \phi''(0)$.

The first order derivative:

(Differentiation is done by using the well-known product rule).

$$\phi'(t) = e^t \underline{\xi}(I - e^t Q)^{-1}\underline{1} + (e^t - 1)\frac{d}{dt}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right)$$

At $t = 0$:

$$\phi'(0) = \underline{\xi}(I - Q)^{-1}\underline{1}$$

Therefore we have that $E(N) = \underline{\xi}(I - Q)^{-1}\underline{1}$.

The second order derivative:

(Differentiation is done by using the well-known product rule).

$$\phi''(t) = e^t \underline{\xi}(I - e^t Q)^{-1}\underline{1} + e^t \frac{d}{dt}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right) + e^t \frac{d}{dt}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right) + (e^t - 1)\frac{d^2}{dt^2}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right)$$

At $t = 0$:

$$\phi''(0) = \underline{\xi}(I - Q)^{-1}\underline{1} + \frac{d}{dt}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right)\Big|_{t=0} + \frac{d}{dt}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right)\Big|_{t=0}$$

$$\phi''(0) = \underline{\xi}(I - Q)^{-1}\underline{1} + 2\frac{d}{dt}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right)\Big|_{t=0} \qquad \text{(A23)}$$

Focusing only on the term $\frac{d}{dt}\left(\underline{\xi}(I - e^t Q)^{-1}\underline{1}\right)\Big|_{t=0}$ we obtain:

$$\frac{d}{dt}\left(\underline{\xi}(I-e^{t}Q)^{-1}\underline{1}\right)\Big|_{t=0}$$
$$=\frac{d}{dt}\left(\underline{\xi}\left(\sum_{n=0}^{\infty}e^{nt}Q^{n}\right)\underline{1}\right)\Big|_{t=0}$$
$$=\underline{\xi}\left(\sum_{n=0}^{\infty}ne^{nt}Q^{n}\right)\underline{1}\Big|_{t=0}$$
$$=\underline{\xi}(e^{t}Q+2e^{2t}Q^{2}+3e^{3t}Q^{3}+4e^{4t}Q^{4}+\ldots)\underline{1}\Big|_{t=0}$$
$$=\underline{\xi}(Q+2Q^{2}+3Q^{3}+4Q^{4}+\ldots)\underline{1}$$
$$=\underline{\xi}Q(I+2Q+3Q^{2}+4Q^{3}+\ldots)\underline{1}$$
$$=\underline{\xi}Q(I-Q)^{-2}\underline{1}.$$

By substituting $\frac{d}{dt}\left(\underline{\xi}(I-e^{t}Q)^{-1}\underline{1}\right)\Big|_{t=0}=\underline{\xi}Q(I-Q)^{-2}\underline{1}$ into expression (A23) we obtain

$$\begin{aligned}\phi''(0)&=\underline{\xi}(I-Q)^{-1}\underline{1}+2\underline{\xi}Q(I-Q)^{-2}\underline{1}\\&=\underline{\xi}\left((I-Q)^{-1}+2Q(I-Q)^{-2}\right)\underline{1}\\&=\underline{\xi}\left((I-Q)^{-1}+2Q(I-Q)^{-1}(I-Q)^{-1}\right)\underline{1}\\&=\underline{\xi}\left((I-Q)^{-1}\left(I+2Q(I-Q)^{-1}\right)\right)\underline{1}\qquad\text{(A24)}\\&=\underline{\xi}\left((I-Q)^{-1}(I+Q)(I-Q)^{-1}\right)\underline{1}\qquad\text{(A25)}\\&=\underline{\xi}(I+Q)(I-Q)^{-2}\underline{1}.\end{aligned}$$

Therefore we have that

$$\boxed{E\left(N^{2}\right)=\underline{\xi}(I+Q)(I-Q)^{-2}\underline{1}}$$

To get from expression (A24) to expression (A25) we used the following expansion

$$(I+Q)(I-Q)^{-1}$$
$$=(I+Q)(I+Q+Q^{2}+Q^{3}+\ldots)$$
$$=I+Q+Q^{2}+Q^{3}+\ldots+Q+Q^{2}+Q^{3}+\ldots$$
$$=I+2Q+2Q^{2}+2Q^{3}+2Q^{4}+\ldots$$
$$=I+2Q(I+Q+Q^{2}+Q^{3}+\ldots)$$
$$=I+2Q(I-Q)^{-1}.$$

Probability generating function

(See page 73 of Fu and Lou (2003))

If N is a discrete random variable taking values of non-negative integers $\{0,1,2,...\}$, then the pgf of N is defined as:

$$\varphi(x) = E\left(x^n\right) = \sum_{n=0}^{\infty} x^n P(N=n) \tag{A26}$$

The pgf is given by $\varphi(t) = 1 + (t-1)\underline{\xi}(I - tQ)^{-1}\underline{1}$. This is obtained by using the definition of the pgf given in expression (A26).

$$\begin{aligned}
\varphi(t) &= \sum_{n=1}^{\infty} t^n P(N=n) \\
&= \sum_{n=1}^{\infty} t^n \left(P(N > n-1) - P(N > n)\right) \\
&= \sum_{n=1}^{\infty} t^n \underline{\xi} Q^{n-1}\underline{1} - \sum_{n=1}^{\infty} t^n \underline{\xi} Q^n \underline{1} \\
&= t\sum_{n=1}^{\infty} t^{n-1} \underline{\xi} Q^{n-1}\underline{1} - \sum_{n=1}^{\infty} t^n \underline{\xi} Q^n \underline{1} \\
&= t\sum_{n=0}^{\infty} t^n \underline{\xi} Q^n \underline{1} - \sum_{n=0}^{\infty} t^n \underline{\xi} Q^n \underline{1} + 1
\end{aligned}$$

The scalar 1 is obtained from the fact that $t^0 \underline{\xi} Q^0 \underline{1} = 1$. By factorizing we obtain

$$\begin{aligned}
\varphi(t) &= 1 + (t-1)\left(\sum_{n=0}^{\infty} t^n \underline{\xi} Q^n \underline{1}\right) \\
&= 1 + (t-1)\underline{\xi}\left(\sum_{n=0}^{\infty} t^n Q^n\right)\underline{1} \\
&= 1 + (t-1)\underline{\xi}(I - tQ)^{-1}\underline{1}.
\end{aligned}$$

It is well-known that if the factorial generating function exists in an interval around $t=1$, then the r^{th} factorial moment is given by $E\left((X)_r\right) = \varphi_X^{(r)}(1) = \left.\frac{d^r}{dt^r}\varphi_X(t)\right|_{t=1}$ where $(X)_r$ is the falling factorial $(x)_r = x(x-1)(x-2)...(x-r+1)$. Therefore, in order to find the *first factorial moment*, we will have to calculate the first order derivative of the pgf in the point $t=1$, that is, $E(N) = \varphi'(1)$. This will give us the *first moment* of the run length variable N. Obtaining the second moment of the run length variable N is more difficult. Firstly, we have to find the *second factorial moment* by calculating the second order derivative of the pgf in the point

$t=1$, that is, $E\big(N(N-1)\big)=\varphi''(1)$. Using the fact that $E\big(N(N-1)\big)=E(N^2)-E(N)$ we can obtain the *second momen*t of the run length variable N.

The first order derivative:

(Differentiation is done by using the well-known product rule).

$$\varphi'(t)=\underline{\xi}(I-tQ)^{-1}\underline{1}+(t-1)\frac{d}{dt}\left(\underline{\xi}(I-tQ)^{-1}\underline{1}\right)$$

At $t=1$:

$$\varphi'(1)=\underline{\xi}(I-Q)^{-1}\underline{1}$$

Therefore we have that

$$E(N)=\underline{\xi}(I-Q)^{-1}\underline{1}. \qquad \text{(A27)}$$

The second order derivative:

(Differentiation is done by using the well-known product rule).

$$\varphi''(t)=\frac{d}{dt}\left(\underline{\xi}(I-tQ)^{-1}\underline{1}\right)+\frac{d}{dt}\left(\underline{\xi}(I-tQ)^{-1}\underline{1}\right)+(t-1)\frac{d^2}{dt^2}\left(\underline{\xi}(I-tQ)^{-1}\underline{1}\right)$$

$$\varphi''(t)=2\frac{d}{dt}\left(\underline{\xi}(I-tQ)^{-1}\underline{1}\right)+(t-1)\frac{d^2}{dt^2}\left(\underline{\xi}(I-tQ)^{-1}\underline{1}\right)$$

At $t=1$:

$$\begin{aligned}
\varphi''(1)&=2\frac{d}{dt}\left(\underline{\xi}(I-tQ)^{-1}\underline{1}\right)\Big|_{t=1}\\
&=2\frac{d}{dt}\left(\underline{\xi}\left(\sum_{n=0}^{\infty}t^nQ^n\right)\underline{1}\right)\Big|_{t=1}\\
&=2\underline{\xi}\left(\sum_{n=0}^{\infty}nt^{n-1}Q^n\right)\underline{1}\Big|_{t=1}\\
&=2\underline{\xi}\left(\sum_{n=0}^{\infty}nQ^n\right)\underline{1}\\
&=2\underline{\xi}\left(0Q^0+1Q^1+2Q^2+3Q^3+\ldots\right)\underline{1}\\
&=2\underline{\xi}Q\left(I+2Q+3Q^2+\ldots\right)\underline{1}\\
&=2\underline{\xi}Q(I-Q)^{-2}\underline{1}.
\end{aligned}$$

Therefore we have that

$$E\big(N(N-1)\big)=2\underline{\xi}Q(I-Q)^{-2}\underline{1}. \qquad \text{(A28)}$$

The second moment is derived by using the fact that

$$E\big(N(N-1)\big) = E(N^2) - E(N) \tag{A29}$$

Expression (A29) can be re-written as

$$E(N^2) = E(N) + E\big(N(N-1)\big) \tag{A30}$$

From (A27) and (A28) we know that $E(N) = \underline{\xi}(I-Q)^{-1}\underline{1}$ and $E\big(N(N-1)\big) = 2\underline{\xi}Q(I-Q)^{-2}\underline{1}$. By substituting this into expression (A30) we obtain $E(N^2) = \underline{\xi}(I-Q)^{-1}\underline{1} + 2\underline{\xi}Q(I-Q)^{-2}\underline{1}$. During the derivation of $E(N^2)$ in the mgf section we have shown that $\underline{\xi}(I-Q)^{-1}\underline{1} + 2\underline{\xi}Q(I-Q)^{-2}\underline{1} = \underline{\xi}(I+Q)(I-Q)^{-2}\underline{1}$. Thus

$$E(N^2) = \underline{\xi}(I+Q)(I-Q)^{-2}\underline{1}. \tag{A31}$$

Finally, though expressions (A27) and (A31) we have $E(N) = \underline{\xi}(I-Q)^{-1}\underline{1}$ and $E(N^2) = \underline{\xi}(I+Q)(I-Q)^{-2}\underline{1}$.

Theorem 3:

In Section 6.1.5 we state that the saddlepoint is the solution to the equation $m(t) = \mu$ where $m(t)$ is the first order derivative of the cumulant generating function (cgf) denoted by $\kappa(t)$. Therefore, the saddlepoint is the solution to the equation $\kappa'(t) = \mu$.

Proof to Theorem 3:

For the development of saddlepoint methodology see Daniels (1954) for details on density approximations, Lugannani and Rice (1980) and Daniels (1987) for discussions on tail area approximations and Reid (1988) for a review on saddlepoint techniques. Saddlepoint approximations are constructed by performing various operations on the moment generating function (mgf) and the cgf.

Let X be a random variable with a density function denoted by $f(x)$. Let $\phi(t)$ denote the mgf which is defined as $\phi(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$. The cgf is just the logarithm of the mgf , i.e. $\kappa(t) = \log(\phi(t))$. From $\phi(t)$ we can obtain $f(x)$ by using the Fourier inversion formula as follows

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(it)e^{-itx}dt = \frac{1}{2\pi}\int_{-i\infty}^{i\infty} e^{\kappa(t)-tx}dt \qquad \text{(A32)}$$

where $i = \sqrt{-1}$.

By differentiating the integral in (A32) and setting the result equal to zero we obtain

$$\kappa'(t) = x\,. \qquad \text{(A33)}$$

The solution to (A33) is called the saddlepoint and denoted by $\hat{t}$.

Daniels (1954) used the exponential power series expansion to estimate the integral in (A32) and derived the following approximation for $f(x)$

$$f(x) \approx \left(\frac{1}{2\pi\kappa''(\hat{t})}\right)^{\frac{1}{2}} e^{\kappa(\hat{t})-\hat{t}x} \qquad \text{(A34)}$$

Expression (A34) is referred to as the first-order saddlepoint density approximation where $\hat{t}$ is the unique solution to the saddlepoint equation $\kappa'(t) = x$.

For a rigorous account of the underlying mathematical theory of saddlepoint methods, interested readers can refer to Jensen (1995).

Appendix B: Computer programs

Mathcad program 1:

This program calculates the *ARL* and the probability of a signal for the upper one-sided Shewhart sign chart. . These values are shown in Tables 2.5, 2.6 and 2.7 for *n* = 5, 10 and 15, respectively.

$$\text{Probsignalupper}(n,p,a) := \left[\sum_{i=n-a}^{n} \text{combin}(n,i)\cdot p^{i}\cdot(1-p)^{n-i}\right]$$

$$\text{ARLupper}(n,p,a) := \frac{1}{\text{Probsignalupper}(n,p,a)}$$

$n := 10$

$q := 0.1, 0.2 .. 0.9$

$a := 0, 1 .. n$

$\text{Msignal}_{q\cdot 10,\,a} := \text{Probsignalupper}(n,q,a)$

$\text{Mupper}_{q\cdot 10,\,a} := \text{ARLupper}(n,q,a)$

Take note: The output is given in Tables 2.5, 2.6 and 2.7, respectively.

Mathcad program 2:

This program calculates the *ARL* and the probability of a signal for the lower one-sided Shewhart sign chart. These values are shown in Tables 2.5, 2.6 and 2.7 for *n* = 5, 10 and 15, respectively.

$$Probsig(n,p,a) := \sum_{i=0}^{a} combin(n,i)\cdot p^{i}\cdot(1-p)^{n-i}$$

$$ARLlower(n,p,a) := \frac{1}{Probsig(n,p,a)}$$

$q := 0.1, 0.2 .. 0.9$

$n := 10$

$a := 0, 1 .. n$

$Msig_{q\cdot 10, a} := Probsig(n,q,a)$

$Mlower_{q\cdot 10, a} := ARLlower(n,q,a)$

Take note: The output is given in Tables 2.5, 2.6 and 2.7, respectively.

Mathcad program 3:

This program calculates the *ARL's* for the upper- and lower one-sided and two-sided Shewhart sign charts with both warning and action limits.

Upper one-sided chart:

$$p0upper(n,p,w) := \sum_{i=0}^{(w+n-2)/2} combin(n,i)\cdot p^{i}(1-p)^{n-i}$$

$$p1upper(n,p,w,a) := \left[\sum_{i=0}^{(a+n-2)/2} combin(n,i)\cdot p^{i}\cdot(1-p)^{n-i}\right] - \left[\sum_{i=0}^{(w+n-2)/2} combin(n,i)\cdot p^{i}\cdot(1-p)^{n-i}\right]$$

$$ARLupper(n,p,w,a,r) := \frac{1-p1upper(n,p,w,a)^{r}}{\left[1-p1upper(n,p,w,a)-p0upper(n,p,w)\cdot\left(1-p1upper(n,p,w,a)^{r}\right)\right]}$$

$p := 0.5 \quad n := 10$

$a := n\cdot p .. n$

$w := n\cdot p .. n$

$RU1_{a,w} := ARLupper(n,p,w,a,1)$

Take note: The output is given in Table 2.10.

Lower one-sided chart:

$$p0lower(n,p,w) := 1 - \sum_{i=0}^{(n-w)/2} combin(n,i)\cdot p^{i}\cdot(1-p)^{n-i}$$

$$p1lower(n,p,w,a) := \left[\sum_{i=0}^{(n-w)/2} combin(n,i)\cdot p^{i}\cdot(1-p)^{n-i}\right] - \left[\sum_{i=0}^{(n-a)/2} combin(n,i)\cdot p^{i}(1-p)^{n-i}\right]$$

$$ARLlower(n,p,w,a,r) := \frac{\left(1-p1lower(n,p,w,a)^{r}\right)}{1-p1lower(n,p,w,a)-p0lower(n,p,w)\cdot\left(1-p1lower(n,p,w,a)^{r}\right)}$$

$RL1_{a,w} := ARLlower(n,p,w,a,1)$

Take note: The output is given in Table 2.11.

Two-sided chart:

$$ARLtwo(n,p,w,a,r) := \frac{ARLlower(n,p,w,a,r)\cdot ARLupper(n,p,w,a,r)}{ARLlower(n,p,w,a,r)+ARLupper(n,p,w,a,r)}$$

$R1_{a,w} := ARLtwo(n,p,w,a,1)$

Take note: The output is given in Table 2.12.

SAS program 1:

This program calculates the SN_i and T_i statistics shown in Table 2.3.

```
proc iml;

* Number of samples;
nn=15;

* Sample size;
n=5;

signmatrix=j(nn,n,0);
timatrix=j(nn,n,0);

* The known median;
tv = 74;

* The matrix containing the Montgomery (2001) Table 5.2 piston ring data;
matrix = {
74.012 74.015 74.030 73.986 74.000, 73.995 74.010 73.990 74.015 74.001,
73.987 73.999 73.985 74.000 73.990, 74.008 74.010 74.003 73.991 74.006,
74.003 74.000 74.001 73.986 73.997, 73.994 74.003 74.015 74.020 74.004,
74.008 74.002 74.018 73.995 74.005, 74.001 74.004 73.990 73.996 73.998,
74.015 74.000 74.016 74.025 74.000, 74.030 74.005 74.000 74.016 74.012,
74.001 73.990 73.995 74.010 74.024, 74.015 74.020 74.024 74.005 74.019,
74.035 74.010 74.012 74.015 74.026, 74.017 74.013 74.036 74.025 74.026,
74.010 74.005 74.029 74.000 74.020};

* Calculating the SNi statistics;
do k = 1 to nn;
      do l = 1 to n;
            if matrix[k,l]>tv then signmatrix[k,l]=1;
            else if matrix[k,l]<tv then signmatrix[k,l]=-1;
            else signmatrix[k,l]=0;
            end;
end;

signvec=signmatrix[,+];

* Calculating the Ti statistics;
do k = 1 to nn;
      do l = 1 to n;
            if matrix[k,l]>tv then timatrix[k,l]=1;
            else timatrix[k,l]=0;
      end;
end;

tivec=timatrix[,+];
si_ti=signvec||tivec;
create newdata from si_ti[colname = {"Si" "Ti"}];
append from si_ti;

proc print data=newdata;
run;
```

SAS program 2:

This program calculates the *ARL*, *SDRL*, 5th, 25th, 50th, 75th and 95th percentile values for the upper one-sided CUSUM sign chart.

Take note: The programs for the lower one-sided and two-sided CUSUM sign charts are omitted, since they are very similar to this program and can easily be obtained by making minor alterations to this program.

```
proc iml;

* For n even, the reference value k is taken to be even;
* For n odd, the reference value k is taken to be odd;
* Restriction h <= n-k;

* The reference value;
k=1;

* The decision interval;
h=4;

* The sample size;
n=5;

* The z-value will be used in the calculation of the pmf, P(N=z), and the cdf,
P(N<=z);
z=10000;

* The following values will be used to calculate the 5th, 25th, 50th, 75th and
95th percentiles, respectively;
p5p=0.05;
p25p=0.25;
p50p=0.5;
p75p=0.75;
p95p=0.95;

* Calculating the state space;
SRn =do(-n,n,2)`;
S=j(nrow(SRn),1,1);

do i = 1 to nrow(SRn);
      S[i,]=min(h,(max(0,SRn[i,]-k)));
end;

do i = 1 to nrow(S);
      do j = 1 to nrow(S);
            if i=j then S[i,]=S[i,];
            else if S[i,]=S[j,] then S[j,]=999;
      end;
end;

S=S[loc(S<999)];

* Defining the vector eta used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
```

```
eta=j(1,nrow(S)-1,1);

do i = 1 to nrow(S)-1;
      if i = 1 then eta[1,i]=1;
      else eta[1,i]=0;
end;

* Calculating the transition probability matrix;
P = j( nrow(S) , nrow(S), 0);
T = j(n+1,1,1);

do x = 0 to n;
      T[x+1,]=pdf('BINOM',x,0.5,n);
end;

* Calculating the first column of the transition matrix;
do i = 1 to nrow(S)-1;
      small_t = (k - S[i,] + n)/2;
      P[i,1] = sum(T[1:(small_t + 1),]);
end;

* Calculating the middle columns of the transition matrix;
do j = 2 to nrow(S)-1;
      do i = 1 to nrow(S)-1;
            small_t=ceil((S[j,] + k - S[i,] + n) / 2);
            P[i,j]=T[small_t+1,];
      end;
end;

* Calculating the last column of the transition matrix;
do i = 1 to nrow(S)-1;
      P[i,nrow(S)] = 1 - sum ( P[i, 1:(nrow(S)-1)] );
end;

P[nrow(S),nrow(S)]=1;

* Defining the vector one used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
one = j(nrow(S)-1,1,1);

* Defining the matrix Q used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
Q = P[1:nrow(S)-1,1:nrow(S)-1];

* Defining the identity matrix I used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
identity = I(nrow(S)-1);

* Calculating the 5th, 25th, 50th, 75th and 95th percentiles;
pmf=j(z,1,1);
cdf=j(z,1,1);
cdf_5th_p=j(z,1,1);
cdf_25th_p=j(z,1,1);
cdf_50th_p=j(z,1,1);
cdf_75th_p=j(z,1,1);
cdf_95th_p=j(z,1,1);
```

```
do i = 1 to z;
      pmf[i,1] = eta * (Q**(i-1)) * (identity - Q) * one;
      cdf[i,1]=sum(pmf[1:i,1]);
end;

index=j(z,1,1);

do i = 2 to z;
      index[i,]=index[i-1,]+1;
end;

* Calculating the 5th percentile;
do i = 1 to z;
      cdf_5th_p[i,]=cdf[i,];
      if cdf_5th_p[i,]>=p5p then cdf_5th_p[i,]=999;
end;

cdf_5th_p=cdf_5th_p[loc(cdf_5th_p<999)];
if cdf_5th_p[1,]=999 then percentile_p5p=1;
else percentile_p5p=nrow(cdf_5th_p)+1;

* Calculating the 25th percentile;
do i = 1 to z;
      cdf_25th_p[i,]=cdf[i,];
      if cdf_25th_p[i,]>=p25p then cdf_25th_p[i,]=999;
end;

cdf_25th_p=cdf_25th_p[loc(cdf_25th_p<999)];
if cdf_25th_p[1,]=999 then percentile_p25p=1;
else percentile_p25p=nrow(cdf_25th_p)+1;

* Calculating the 50th percentile;
do i = 1 to z;
      cdf_50th_p[i,]=cdf[i,];
      if cdf_50th_p[i,]>=p50p then cdf_50th_p[i,]=999;
end;

cdf_50th_p=cdf_50th_p[loc(cdf_50th_p<999)];
if cdf_50th_p[1,]=999 then percentile_p50p=1;
else percentile_p50p=nrow(cdf_50th_p)+1;

* Calculating the 75th percentile;
do i = 1 to z;
      cdf_75th_p[i,]=cdf[i,];
      if cdf_75th_p[i,]>=p75p then cdf_75th_p[i,]=999;
end;

cdf_75th_p=cdf_75th_p[loc(cdf_75th_p<999)];
if cdf_75th_p[1,]=999 then percentile_p75p=1;
else percentile_p75p=nrow(cdf_75th_p)+1;

* Calculating the 95th percentile;
do i = 1 to z;
      cdf_95th_p[i,]=cdf[i,];
      if cdf_95th_p[i,]>=p95p then cdf_95th_p[i,]=999;
end;
```

```
cdf_95th_p=cdf_95th_p[loc(cdf_95th_p<999)];
if cdf_95th_p[1,]=999 then percentile_p95p=1;
else percentile_p95p=nrow(cdf_95th_p)+1;

* Calculating the average run length (ARL);
ARL = eta * inv(identity-Q) * one;

* Calculating the second moment;
N2 = eta * (identity + Q) * (inv((identity-Q)**2)) * one;

* Calculating the standard deviation;
SDRL = sqrt (N2 - ((ARL)**2) );

* Calculating the two-sided ARL;
ARL_two=(ARL*ARL)/(ARL+ARL);

* Printing the output;
print_cdf=index||cdf;
print_pmf=index||pmf;

print   k [label='Reference value']
        , h [label='Desicion interval']
        , n [label='Sample size']
        , S [label = 'State Space']
        , P [label='Transition probability matrix' format=.3]
        , ARL [label='Average run length' format=.2]
        , ARL_two [label = 'The ARL of the two-sided chart' format=.2]
        , SDRL [label='Standard Deviation of the run length' format=.2]
        , N2 [label='Second moment' format=.2]
        , percentile_p5p [label='Fifth percentile']
        , percentile_p25p [label='25th percentile']
        , percentile_p50p [label='50th percentile']
        , percentile_p75p [label='75th percentile']
        , percentile_p95p [label='95th percentile'];
run;
```

SAS program 3:

This program calculates the sign test statistics (SN_i), T_i-statistics, upper CUSUM statistics (S_i^+) and lower CUSUM statistics (S_i^-) for the Montgomery (2001) piston ring data.

```
proc iml;

* Number of samples;
nn=15;

* Sample size;
n=5;

* The known median;
tv = 74;

signmatrix=j(nn,n,0);
timatrix=j(nn,n,0);

* A matrix containing the Montgomery (2001) Table 5.2 piston ring data;
matrix =
{74.012 74.015 74.030 73.986 74.000,
73.995 74.010 73.990 74.015 74.001,
73.987 73.999 73.985 74.000 73.990,
74.008 74.010 74.003 73.991 74.006,
74.003 74.000 74.001 73.986 73.997,
73.994 74.003 74.015 74.020 74.004,
74.008 74.002 74.018 73.995 74.005,
74.001 74.004 73.990 73.996 73.998,
74.015 74.000 74.016 74.025 74.000,
74.030 74.005 74.000 74.016 74.012,
74.001 73.990 73.995 74.010 74.024,
74.015 74.020 74.024 74.005 74.019,
74.035 74.010 74.012 74.015 74.026,
74.017 74.013 74.036 74.025 74.026,
74.010 74.005 74.029 74.000 74.020};

* Calculating the sign test statistics, SNi;
do k = 1 to nn;
      do l = 1 to n;
            if matrix[k,l]>tv then signmatrix[k,l]=1;
            else if matrix[k,l]<tv then signmatrix[k,l]=-1;
            else signmatrix[k,l]=0;
            end;
end;

signvec=signmatrix[,+];

* Calculating the Ti statistics;
do k = 1 to nn;
      do l = 1 to n;
      if matrix[k,l]>tv then timatrix[k,l]=1;
      else timatrix[k,l]=0;
      end;
end;
```

```
tivec=timatrix[,+];

* Calculating the CUSUM statistics;

* Specifying the reference value;
k=2;

* The starting values are set equal to zero;
SNpluszero=0;
SNminuszero=0;

SNplus=j(nn,1,0);
SNminus=j(nn,1,0);

do l = 2 to nn;
      SNplus[1,]=max(0,SNpluszero+(signvec[1,])-k);
      SNminus[1,]=min(0,SNminuszero+(signvec[1,])+k);
      SNplus[l,]=max(0,SNplus[l-1,]+(signvec[l,])-k);
      SNminus[l,]=min(0,SNminus[l-1,]+(signvec[l,])+k);
end;

Nplus=j(nn,1,0);
Nminus=j(nn,1,0);

do l = 1 to nn;
      if SNplus[l,]=0 then Nplus[l,]=0;
      else Nplus[l,]=Nplus[l-1,]+1;
end;

do l = 1 to nn;
      if SNminus[l,]=0 then Nminus[l,]=0;
      else Nminus[l,]=Nminus[l-1,]+1;
end;

* Printing the output;
print signvec [label='The SNi statistics'],
      tivec [label='The Ti statistics'],
      SNplus [label='The upper CUSUM statistics'],
      SNminus [label='The lower CUSUM statistics'];
```

SAS program 4:

This program calculates the *ARL*, *SDRL*, 5th, 25th, 50th, 75th and 95th percentile values for the EWMA sign chart.

```
proc iml;

* Number of subintervals between UCL and LCL;
NN=5;

* Sample size;
n=6;

* p=0.5 when the process is in-control;
p=0.5;

* The EWMA parameter: the multiplier;
L=2;

* The EWMA parameter: the smoothing constant;
lambda=1;

* The z-value will be used in the calculation of the percentiles;
z=10000;

* Calculating the control limits;
UCL = L * 2 * sqrt(n*p*(1-p)) * sqrt(lambda/(2-lambda));
LCL = -UCL;
S=j(NN,1,0);

* The interval between the UCL and LCL are divided into subintervals of width
2*delta;
delta = ((UCL-LCL)/NN)/2;
S[1,1] = UCL - delta;

do i = 2 to NN;
      S[i,1]=S[i-1,1]-2*delta;
end;

Q_a=j(NN,NN,0);
Q_b=j(NN,NN,0);
Q=j(NN,NN,0);

do i = 1 to NN;
      do j = 1 to NN;
      Q_a[i,j]=floor(((((S[j,]-delta) - (1-lambda)*S[i,])/lambda) + n)/2);
      end;
end;

do i = 1 to NN;
      do j = 1 to NN;
      Q_b[i,j]=floor(((((S[j,]+delta) - (1-lambda)*S[i,])/lambda) + n)/2);
      end;
end;

do i = 1 to NN;
```

```
      do j = 1 to NN;
      Q[i,j]=cdf('BINOMIAL',(Q_b[i,j]),p,n)-cdf('BINOMIAL',(Q_a[i,j]),p,n);
      end;
end;

eta=j(1,NN,1);
do i = 1 to NN;
      if i = 1 then eta[1,i]=1;
      else eta[1,i]=0;
end;

* Defining the vector one used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
one=j(NN,1,1);

* Defining the identity matrix I used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
identity = I(NN);

* Calculating the 5th, 25th, 50th, 75th and 95th percentiles;
p5p=0.05;
p25p=0.25;
p50p=0.5;
p75p=0.75;
p95p=0.95;
pmf=j(z,1,1);
cdf=j(z,1,1);
cdf_5th_p=j(z,1,1);
cdf_25th_p=j(z,1,1);
cdf_50th_p=j(z,1,1);
cdf_75th_p=j(z,1,1);
cdf_95th_p=j(z,1,1);

do i = 1 to z;
      pmf[i,1] = eta * (Q**(i-1)) * (identity - Q) * one;
      cdf[i,1]=sum(pmf[1:i,1]);
end;

index=j(z,1,1);

do i = 2 to z;
      index[i,]=index[i-1,]+1;
end;

* Calculating the 5th percentile;
do i = 1 to z;
      cdf_5th_p[i,]=cdf[i,];
      if cdf_5th_p[i,]>=p5p then cdf_5th_p[i,]=999;
end;
cdf_5th_p=cdf_5th_p[loc(cdf_5th_p<999)];
if cdf_5th_p[1,]=999 then percentile_p5p=1;
else percentile_p5p=nrow(cdf_5th_p)+1;

* Calculating the 25th percentile;
do i = 1 to z;
      cdf_25th_p[i,]=cdf[i,];
      if cdf_25th_p[i,]>=p25p then cdf_25th_p[i,]=999;
```

```
end;
cdf_25th_p=cdf_25th_p[loc(cdf_25th_p<999)];
if cdf_25th_p[1,]=999 then percentile_p25p=1;
else percentile_p25p=nrow(cdf_25th_p)+1;

* Calculating the 50th percentile;
do i = 1 to z;
      cdf_50th_p[i,]=cdf[i,];
      if cdf_50th_p[i,]>=p50p then cdf_50th_p[i,]=999;
end;
cdf_50th_p=cdf_50th_p[loc(cdf_50th_p<999)];
if cdf_50th_p[1,]=999 then percentile_p50p=1;
else percentile_p50p=nrow(cdf_50th_p)+1;

* Calculating the 75th percentile;
do i = 1 to z;
      cdf_75th_p[i,]=cdf[i,];
      if cdf_75th_p[i,]>=p75p then cdf_75th_p[i,]=999;
end;
cdf_75th_p=cdf_75th_p[loc(cdf_75th_p<999)];
if cdf_75th_p[1,]=999 then percentile_p75p=1;
else percentile_p75p=nrow(cdf_75th_p)+1;

* Calculating the 95th percentile;
do i = 1 to z;
      cdf_95th_p[i,]=cdf[i,];
      if cdf_95th_p[i,]>=p95p then cdf_95th_p[i,]=999;
end;
cdf_95th_p=cdf_95th_p[loc(cdf_95th_p<999)];
if cdf_95th_p[1,]=999 then percentile_p95p=1;
else percentile_p95p=nrow(cdf_95th_p)+1;

* Calculating the average run length (ARL);
ARL = eta*ginv(identity-Q)*one;

* Calculating the second moment;
N2 = eta * (identity + Q) * (ginv((identity-Q)**2)) * one;

* Calculating the standard deviation;
SDRL = sqrt (N2 - ((ARL)**2) );

* Printing the output;
print_cdf=index||cdf;
print_pmf=index||pmf;

print       UCL [label='Upper control limit'],
            LCL [label='Lower control limit'],
            delta,
            Q,
            ARL [label = 'Average run length' format=.2]
            SDRL [label = 'Standard deviation of the run length' format=.2],
            percentile_p5p [label='Fifth percentile'],
            percentile_p25p [label='25th percentile'],
            percentile_p50p [label='50th percentile'],
            percentile_p75p [label='75th percentile'],
            percentile_p95p [label='95th percentile'];
```

SAS program 5:

This program calculates the signed-rank (SR_i) statistics for the Shewhart-type signed-rank chart using the Montgomery (2001) piston ring data.

```
proc iml;

* Number of samples;
nn=15;

* Sample size;
n=5;

wsrmatrix = j(nn,n,0);
sgnmatrix = j(nn,n,0);
rank_abs_diff = j(nn,n,0);
final_rank_abs_diff=j(nn,n,0);

* The known median;
tv=74;

tvmatrix = j(nn,n,tv);

* A matrix containing the Montgomery (2001) Table 5.2 piston ring data;
obs = {
74.012 74.015 74.030 73.986 74.000,
73.995 74.010 73.990 74.015 74.001,
73.987 73.999 73.985 74.000 73.990,
74.008 74.010 74.003 73.991 74.006,
74.003 74.000 74.001 73.986 73.997,
73.994 74.003 74.015 74.020 74.004,
74.008 74.002 74.018 73.995 74.005,
74.001 74.004 73.990 73.996 73.998,
74.015 74.000 74.016 74.025 74.000,
74.030 74.005 74.000 74.016 74.012,
74.001 73.990 73.995 74.010 74.024,
74.015 74.020 74.024 74.005 74.019,
74.035 74.010 74.012 74.015 74.026,
74.017 74.013 74.036 74.025 74.026,
74.010 74.005 74.029 74.000 74.020};

* Calculating the SRi statistics;
diff=obs-tvmatrix;

do k = 1 to nn;
      do l = 1 to n;
      if diff[k,l]>0 then sgnmatrix[k,l]=1;
      else if diff[k,l]<0 then sgnmatrix[k,l]=-1;
      else if diff[k,l]=0 then sgnmatrix[k,l]=0;
      end;
end;

abs_diff=abs(diff);

do i = 1 to nn;
      rank_abs_diff[i,]=rank(abs_diff[i,]);
```

```
end;

do i = 1 to nn;
      do j = 1 to n;
            do k = 1 to n;
            if abs_diff[i,j]=abs_diff[i,k] then

      final_rank_abs_diff[i,j]=(rank_abs_diff[i,j]+rank_abs_diff[i,k])/2;
            end;
      end;
end;

do i = 1 to nn;
      do j = 1 to n;
            do k = 1 to n;
            if abs_diff[i,j]=abs_diff[i,k] then
            final_rank_abs_diff[i,k]=final_rank_abs_diff[i,j];
            end;
      end;
end;


do k = 1 to nn;
      do l = 1 to n;
            wsrmatrix[k,l]=sgnmatrix[k,l]*final_rank_abs_diff[k,l];
      end;
end;

SRi = wsrmatrix[,+];

* Printing the output;
print SRi [label='The SRi statistics'];
```

SAS program 6:

This program calculates the $FAR's$ and $ARL_0's$ for the two-sided Shewhart signed-rank chart and the two-sided Shewhart signed-rank-like chart, respectively.

```
proc iml;

*Number of simulations;
numsim=10000;

ARLmatrix=j(numsim,1,0);

* Starting the simulation study;
do ss = 1 to numsim;

* The upper control limit;
ucl = 10;
ucl = ucl-1;

* The lower control limit;
lcl = -ucl;

* Number of samples (must theoretically go to infinity;
nn=10000;

* Sample size;
n=10;

wsrmatrix = j(nn,n,0);
sgnmatrix = j(nn,n,0);
rank_abs_diff = j(nn,n,0);
final_rank_abs_diff=j(nn,n,0);

* The median of the standard normal distribution;
tv=0;

tvmatrix = j(nn,n,tv);
obs = j(nn,n,0);

* Genrating oservations from a standard normal distribution;
call randgen(obs,'normal');

diff=obs-tvmatrix;

do k = 1 to nn;
      do l = 1 to n;
            if diff[k,l]>0 then sgnmatrix[k,l]=1;
            else if diff[k,l]<0 then sgnmatrix[k,l]=-1;
            else if diff[k,l]=0 then sgnmatrix[k,l]=0;
      end;
end;

abs_diff=abs(diff);

do i = 1 to nn;
```

```
      rank_abs_diff[i,]=rank(abs_diff[i,]);
end;

do i = 1 to nn;
      do j = 1 to n;
            do k = 1 to n;
                  if abs_diff[i,j]=abs_diff[i,k] then

      final_rank_abs_diff[i,j]=(rank_abs_diff[i,j]+rank_abs_diff[i,k])/2;
            end;
      end;
end;

do i = 1 to nn;
      do j = 1 to n;
            do k = 1 to n;
                  if abs_diff[i,j]=abs_diff[i,k] then
                  final_rank_abs_diff[i,k]=final_rank_abs_diff[i,j];
            end;
      end;
end;


do k = 1 to nn;
      do l = 1 to n;
            wsrmatrix[k,l]=sgnmatrix[k,l]*final_rank_abs_diff[k,l];
      end;
end;

SRi = wsrmatrix[,+];
count = j(nn,1,0);

do l = 1 to nn;
      if SRi[l,]>=ucl then count[l,]=999;
      if SRi[l,]<=lcl then count[l,]=999;
end;

do ll = 1 to nn;
      if count[ll,]=999 then goto skip;
end;

skip: ARL = ll;
ARLmatrix[ss,1]=ARL;
end;

* The simulated average run length (ARL);
simulatedARL = ARLmatrix[+,]/nrow(ARLmatrix);

* The simulated false alarm rate (FAR);
FAR = 1/simulatedARL;

ucl = ucl+1;

print       n [label='Sample size'],
            ucl [label = 'Upper control limit'],
            simulatedARL [label = 'ARL' format=.3],
            FAR [label = 'FAR' format=.3];
```

SAS program 7:

This program calculates the *ARL*, *SDRL*, 5th, 25th, 50th, 75th and 95th percentile values for the upper one-sided CUSUM signed-rank chart.

Take note: The programs for the lower one-sided and two-sided CUSUM signed-rank charts are omitted, since they are very similar to this program and can easily be obtained by making minor alterations to this program.

```
proc iml;

* The reference value;
k=2;

* The decision interval;
h=8;

* The sample size;
n=4;

* The z-value will be used in the calculation of the pmf, P(N=z), and the cdf,
P(N<=z);
z=10000;

* The following values will be used to calculate the 5th, 25th, 50th, 75th and
95th percentiles, respectively;
p5p=0.05;
p25p=0.25;
p50p=0.5;
p75p=0.75;
p95p=0.95;

* Calculating the state space;
SRn =do((-n*(n+1)/2),(n*(n+1)/2),2)`;
S=j(nrow(SRn),1,1);

do i = 1 to nrow(SRn);
      S[i,]=min(h,(max(0,SRn[i,]-k)));
end;

do i = 1 to nrow(S);
      do j = 1 to nrow(S);
            if i=j then S[i,]=S[i,];
            else if S[i,]=S[j,] then S[j,]=999;
      end;
end;

S=S[loc(S<999)];

if h > (SRn[nrow(SRn),]-k) then print "Not possible";

* Defining the vector eta used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
eta=j(1,nrow(S)-1,1);
```

```
do i = 1 to nrow(S)-1;
      if i = 1 then eta[1,i]=1;
      else eta[1,i]=0;
end;

* Calculating the transition probability matrix;
P = j( nrow(S) , nrow(S), 0);

* Wilcoxon siged-rank probabilities for a sample size of 4;
if n = 4 then do;
      T4 = j((n*(n+1)/2)+1,1,1);
      T4[1:3,]=1;
      T4[4:8,]=2;
      T4[9:11,]=1;
      T = T4/(2**n);
end;

* Wilcoxon siged-rank probabilities for a sample size of 5;
if n = 5 then do;
      T5 = j((n*(n+1)/2)+1,1,1);
      T5[1:3,]=1;
      T5[4:5,]=2;
      T5[6:11,]=3;
      T5[12:13,]=2;
      T5[14:16,]=1;
      T = T5/(2**n);
end;

* Wilcoxon siged-rank probabilities for a sample size of 6;
if n = 6 then do;
      T6 = j((n*(n+1)/2)+1,1,1);
      T6[1:3,]=1;
      T6[4:5,]=2;
      T6[6,]=3;
      T6[7:9,]=4;
      T6[10:13,]=5;
      T6[14:16,]=4;
      T6[17,]=3;
      T6[18:19,]=2;
      T6[20:22,]=1;
      T = T6/(2**n);
end;


* Wilcoxon siged-rank probabilities for a sample size of 7;
if n = 7 then do;
      T7 = j((n*(n+1)/2)+1,1,1);
      T7[1:3,]=1;
      T7[4:5,]=2;
      T7[6,]=3;
      T7[7,]=4;
      T7[8:9,]=5;
      T7[10,]=6;
      T7[11:12,]=7;
      T7[13:17,]=8;
      T7[18:19,]=7;
      T7[20,]=6;
```

```
      T7[21:22,]=5;
      T7[23,]=4;
      T7[24,]=3;
      T7[25:26,]=2;
      T7[27:29,]=1;
      T = T7/(2**n);
end;

* Wilcoxon siged-rank probabilities for a sample size of 8;
if n = 8 then do;
      T8 = j((n*(n+1)/2)+1,1,1);
      T8[1:3,]=1;
      T8[4:5,]=2;
      T8[6,]=3;
      T8[7,]=4;
      T8[8,]=5;
      T8[9,]=6;
      T8[10,]=7;
      T8[11,]=8;
      T8[12,]=9;
      T8[13,]=10;
      T8[14,]=11;
      T8[15,]=12;
      T8[16:18,]=13;
      T8[19,]=14;
      T8[20:22,]=13;
      T8[23,]=12;
      T8[24,]=11;
      T8[25,]=10;
      T8[26,]=9;
      T8[27,]=8;
      T8[28,]=7;
      T8[29,]=6;
      T8[30,]=5;
      T8[31,]=4;
      T8[32,]=3;
      T8[33:34,]=2;
      T8[35:37,]=1;
      T = T8/(2**n);
end;

* Wilcoxon siged-rank probabilities for a sample size of 9;
if n = 9 then do;
      T9 = j((n*(n+1)/2)+1,1,1);
      T9[1:3,]=1;
      T9[4:5,]=2;
      T9[6,]=3;
      T9[7,]=4;
      T9[8,]=5;
      T9[9,]=6;
      T9[10,]=8;
      T9[11,]=9;
      T9[12,]=10;
      T9[13,]=12;
      T9[14,]=13;
      T9[15,]=15;
      T9[16,]=17;
```

```
        T9[17,]=18;
        T9[18,]=19;
        T9[19:20,]=21;
        T9[21,]=22;
        T9[22:25,]=23;
        T9[26,]=22;
        T9[27:28,]=21;
        T9[29,]=19;
        T9[30,]=18;
        T9[31,]=17;
        T9[32,]=15;
        T9[33,]=13;
        T9[34,]=12;
        T9[35,]=10;
        T9[36,]=9;
        T9[37,]=8;
        T9[38,]=6;
        T9[39,]=5;
        T9[40,]=4;
        T9[41,]=3;
        T9[42:43,]=2;
        T9[44:46,]=1;
        T = T9/(2**n);
end;

* Wilcoxon siged-rank probabilities for a sample size of 10;
if n = 10 then do;
        T10 = j((n*(n+1)/2)+1,1,1);
        T10[1:3,]=1;
        T10[4:5,]=2;
        T10[6,]=3;
        T10[7,]=4;
        T10[8,]=5;
        T10[9,]=6;
        T10[10,]=8;
        T10[11,]=10;
        T10[12,]=11;
        T10[13,]=13;
        T10[14,]=15;
        T10[15,]=17;
        T10[16,]=20;
        T10[17,]=22;
        T10[18,]=24;
        T10[19,]=27;
        T10[20,]=29;
        T10[21,]=31;
        T10[22,]=33;
        T10[23,]=35;
        T10[24,]=36;
        T10[25,]=38;
        T10[26:27,]=39;
        T10[28:29,]=40;
        T10[30:31,]=39;
        T10[32,]=38;
        T10[33,]=36;
        T10[34,]=35;
        T10[35,]=33;
```

```
        T10[36,]=31;
        T10[37,]=29;
        T10[38,]=27;
        T10[39,]=24;
        T10[40,]=22;
        T10[41,]=20;
        T10[42,]=17;
        T10[43,]=15;
        T10[44,]=13;
        T10[45,]=11;
        T10[46,]=10;
        T10[47,]=8;
        T10[48,]=6;
        T10[49,]=5;
        T10[50,]=4;
        T10[51,]=3;
        T10[52:53,]=2;
        T10[54:56,]=1;
        T = T10/(2**n);
end;

* Calculating the first column of the transition matrix;
do i = 1 to nrow(S)-1;
        small_t = (k - S[i,] + (n*(n+1)/2) ) / 2;
        P[i,1] = sum(T[1:(small_t + 1),]);
end;

* Calculating the middle columns of the transition matrix;
do j = 2 to nrow(S)-1;
        do i = 1 to nrow(S)-1;
                small_t=(S[j,] + k - S[i,] + (n*(n+1)/2)) / 2;
                P[i,j] = T[small_t + 1,];
        end;
end;

* Calculating the last column of the transition matrix;
do i = 1 to nrow(S)-1;
        small_t=((S[nrow(S),] + k - S[i,] + (n*(n+1)/2)) / 2)-1;
        P[i,nrow(S)] = 1- sum(T[ 1:(small_t + 1),]);
end;

P[nrow(S),nrow(S)]=1;

* Defining the vector one used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
one = j(nrow(S)-1,1,1);

* Defining the matrix Q used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
Q = P[1:nrow(S)-1,1:nrow(S)-1];

* Defining the identity matrix I used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
identity = I(nrow(S)-1);

* Calculating the 5th, 25th, 50th, 75th and 95th percentiles;
pmf=j(z,1,1);
```

```
cdf=j(z,1,1);
cdf_5th_p=j(z,1,1);
cdf_25th_p=j(z,1,1);
cdf_50th_p=j(z,1,1);
cdf_75th_p=j(z,1,1);
cdf_95th_p=j(z,1,1);

do i = 1 to z;
      pmf[i,1] = eta * (Q**(i-1)) * (identity - Q) * one;
      cdf[i,1]=sum(pmf[1:i,1]);
end;

index=j(z,1,1);

do i = 2 to z;
      index[i,]=index[i-1,]+1;
end;

* Calculating the 5th percentile;
do i = 1 to z;
      cdf_5th_p[i,]=cdf[i,];
      if cdf_5th_p[i,]>=p5p then cdf_5th_p[i,]=999;
end;

cdf_5th_p=cdf_5th_p[loc(cdf_5th_p<999)];

if cdf_5th_p[1,]=999 then percentile_p5p=1;
else percentile_p5p=nrow(cdf_5th_p)+1;

* Calculating the 25th percentile;
do i = 1 to z;
      cdf_25th_p[i,]=cdf[i,];
      if cdf_25th_p[i,]>=p25p then cdf_25th_p[i,]=999;
end;

cdf_25th_p=cdf_25th_p[loc(cdf_25th_p<999)];

if cdf_25th_p[1,]=999 then percentile_p25p=1;
else percentile_p25p=nrow(cdf_25th_p)+1;

* Calculating the 50th percentile;
do i = 1 to z;
      cdf_50th_p[i,]=cdf[i,];
      if cdf_50th_p[i,]>=p50p then cdf_50th_p[i,]=999;
end;

cdf_50th_p=cdf_50th_p[loc(cdf_50th_p<999)];

if cdf_50th_p[1,]=999 then percentile_p50p=1;
else percentile_p50p=nrow(cdf_50th_p)+1;

* Calculating the 75th percentile;
do i = 1 to z;
      cdf_75th_p[i,]=cdf[i,];
      if cdf_75th_p[i,]>=p75p then cdf_75th_p[i,]=999;
end;
```

```
cdf_75th_p=cdf_75th_p[loc(cdf_75th_p<999)];

if cdf_75th_p[1,]=999 then percentile_p75p=1;
else percentile_p75p=nrow(cdf_75th_p)+1;

* Calculating the 95th percentile;
do i = 1 to z;
      cdf_95th_p[i,]=cdf[i,];
      if cdf_95th_p[i,]>=p95p then cdf_95th_p[i,]=999;
end;

cdf_95th_p=cdf_95th_p[loc(cdf_95th_p<999)];

if cdf_95th_p[1,]=999 then percentile_p95p=1;
else percentile_p95p=nrow(cdf_95th_p)+1;

* Calculating the average run length (ARL);
ARL = eta * inv(identity-Q) * one;

* Calculating the second moment;
N2 = eta * (identity + Q) * (inv((identity-Q)**2)) * one;

* Calculating the standard deviation;
SDRL = sqrt (N2 - ((ARL)**2) );

* Printing the output;
print_cdf=index||cdf;
print_pmf=index||pmf;

print       k [label='Reference value'],
            h [label='Desicion interval'],
            n [label='Sample size'],
            S [label = 'State Space'],
            P [label='Transition probability matrix' format=fract.],
            ARL [label='ARL' format=.2],
            SDRL [label='SDRL' format=.2],
            percentile_p5p [label='5th'],
            percentile_p25p [label='25th'],
            percentile_p50p [label='50th'],
            percentile_p75p [label='75th'],
            percentile_p95p [label='95th'];
run;
```

SAS program 8:

This program calculates the *ARL*, *SDRL*, 5th, 25th, 50th, 75th and 95th percentile values for the EWMA signed-rank chart.

```
proc iml;

* Number of subintervals between UCL and LCL;
NN=5;

* Sample size;
n=10;

* p=0.5 when the process is in-control;
p=0.5;

* The EWMA parameter: the multiplier;
L=1;

* The EWMA parameter: the smoothing constant;
lambda=0.2;

* The z-value will be used in the calculation of the percentiles;
z=10000;

* The in-control standard devation of the signed-rank statistic;
stdev=sqrt ( n * (n+1) * (2*n+1) / 6);

* Wilcoxon siged-rank probabilities for a sample size of 4;
if n = 4 then do;
      T4 = j((n*(n+1)/2)+1,1,1);
      T4[1:3,]=1;
      T4[4:8,]=2;
      T4[9:11,]=1;
      T = T4/(2**n);
end;

* Wilcoxon signed-rank probabilities for a sample size of 5;
if n = 5 then do;
      T5 = j((n*(n+1)/2)+1,1,1);
      T5[1:3,]=1;
      T5[4:5,]=2;
      T5[6:11,]=3;
      T5[12:13,]=2;
      T5[14:16,]=1;
      T = T5/(2**n);
end;

* Wilcoxon signed-rank probabilities for a sample size of 6;
if n = 6 then do;
      T6 = j((n*(n+1)/2)+1,1,1);
      T6[1:3,]=1;
      T6[4:5,]=2;
      T6[6,]=3;
      T6[7:9,]=4;
```

```
      T6[10:13,]=5;
      T6[14:16,]=4;
      T6[17,]=3;
      T6[18:19,]=2;
      T6[20:22,]=1;
      T = T6/(2**n);
end;


* Wilcoxon signed-rank probabilities for a sample size of 7;
if n = 7 then do;
      T7 = j((n*(n+1)/2)+1,1,1);
      T7[1:3,]=1;
      T7[4:5,]=2;
      T7[6,]=3;
      T7[7,]=4;
      T7[8:9,]=5;
      T7[10,]=6;
      T7[11:12,]=7;
      T7[13:17,]=8;
      T7[18:19,]=7;
      T7[20,]=6;
      T7[21:22,]=5;
      T7[23,]=4;
      T7[24,]=3;
      T7[25:26,]=2;
      T7[27:29,]=1;
      T = T7/(2**n);
end;

* Wilcoxon signed-rank probabilities for a sample size of 8;
if n = 8 then do;
      T8 = j((n*(n+1)/2)+1,1,1);
      T8[1:3,]=1;
      T8[4:5,]=2;
      T8[6,]=3;
      T8[7,]=4;
      T8[8,]=5;
      T8[9,]=6;
      T8[10,]=7;
      T8[11,]=8;
      T8[12,]=9;
      T8[13,]=10;
      T8[14,]=11;
      T8[15,]=12;
      T8[16:18,]=13;
      T8[19,]=14;
      T8[20:22,]=13;
      T8[23,]=12;
      T8[24,]=11;
      T8[25,]=10;
      T8[26,]=9;
      T8[27,]=8;
      T8[28,]=7;
      T8[29,]=6;
      T8[30,]=5;
      T8[31,]=4;
```

```
        T8[32,]=3;
        T8[33:34,]=2;
        T8[35:37,]=1;
        T = T8/(2**n);
end;

* Wilcoxon signed-rank probabilities for a sample size of 9;
if n = 9 then do;
        T9 = j((n*(n+1)/2)+1,1,1);
        T9[1:3,]=1;
        T9[4:5,]=2;
        T9[6,]=3;
        T9[7,]=4;
        T9[8,]=5;
        T9[9,]=6;
        T9[10,]=8;
        T9[11,]=9;
        T9[12,]=10;
        T9[13,]=12;
        T9[14,]=13;
        T9[15,]=15;
        T9[16,]=17;
        T9[17,]=18;
        T9[18,]=19;
        T9[19:20,]=21;
        T9[21,]=22;
        T9[22:25,]=23;
        T9[26,]=22;
        T9[27:28,]=21;
        T9[29,]=19;
        T9[30,]=18;
        T9[31,]=17;
        T9[32,]=15;
        T9[33,]=13;
        T9[34,]=12;
        T9[35,]=10;
        T9[36,]=9;
        T9[37,]=8;
        T9[38,]=6;
        T9[39,]=5;
        T9[40,]=4;
        T9[41,]=3;
        T9[42:43,]=2;
        T9[44:46,]=1;
        T = T9/(2**n);
end;

* Wilcoxon signed-rank probabilities for a sample size of 10;
if n = 10 then do;
        T10 = j((n*(n+1)/2)+1,1,1);
        T10[1:3,]=1;
        T10[4:5,]=2;
        T10[6,]=3;
        T10[7,]=4;
        T10[8,]=5;
        T10[9,]=6;
        T10[10,]=8;
```

```
      T10[11,]=10;
      T10[12,]=11;
      T10[13,]=13;
      T10[14,]=15;
      T10[15,]=17;
      T10[16,]=20;
      T10[17,]=22;
      T10[18,]=24;
      T10[19,]=27;
      T10[20,]=29;
      T10[21,]=31;
      T10[22,]=33;
      T10[23,]=35;
      T10[24,]=36;
      T10[25,]=38;
      T10[26:27,]=39;
      T10[28:29,]=40;
      T10[30:31,]=39;
      T10[32,]=38;
      T10[33,]=36;
      T10[34,]=35;
      T10[35,]=33;
      T10[36,]=31;
      T10[37,]=29;
      T10[38,]=27;
      T10[39,]=24;
      T10[40,]=22;
      T10[41,]=20;
      T10[42,]=17;
      T10[43,]=15;
      T10[44,]=13;
      T10[45,]=11;
      T10[46,]=10;
      T10[47,]=8;
      T10[48,]=6;
      T10[49,]=5;
      T10[50,]=4;
      T10[51,]=3;
      T10[52:53,]=2;
      T10[54:56,]=1;
      T = T10/(2**n);
end;

* Calculating the control limits;
UCL = L * stdev * sqrt(lambda/(2-lambda));
LCL = -UCL;

S=j(NN,1,0);

* The interval between the UCL and LCL are divided into subintervals of width
2*delta;
delta = ((UCL-LCL)/NN)/2;

S[1,1] = UCL - delta;

do i = 2 to NN;
      S[i,1]=S[i-1,1]-2*delta;
```

```
end;

Q_a=j(NN,NN,0);
Q_b=j(NN,NN,0);
Q=j(NN,NN,0);

do i = 1 to NN;
      do j = 1 to NN;
      Q_a[i,j]=floor(((((S[j,]-delta) - (1-lambda)*S[i,])/lambda) +
n*(n+1)/2)/2);
      end;
end;

do i = 1 to NN;
      do j = 1 to NN;
      Q_b[i,j]=floor(((((S[j,]+delta) - (1-lambda)*S[i,])/lambda) +
n*(n+1)/2)/2);
      end;
end;

do i = 1 to NN;
      do j = 1 to NN;
      lower = Q_a[i,j];
      upper = Q_b[i,j];
      if lower < 0 then if lower*upper > 0 then Q[i,j]=0;
      if lower > n*(n+1)/2 then lower_term = 1;
      else if lower < 0 then lower_term = 0;
      else lower_term = sum(T[1:(lower+1),]);
      if upper > n*(n+1)/2 then upper_term = 1;
      else if upper < 0 then upper_term = 0;
      else upper_term = sum(T[1:(upper+1),]);
      Q[i,j] = upper_term - lower_term;
      end;
end;

* Defining the vector eta used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
eta=j(1,NN,1);
do i = 1 to NN;
      if i = 1 then eta[1,i]=1;
      else eta[1,i]=0;
end;

* Defining the vector one used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
one=j(NN,1,1);

* Defining the identity matrix I used in the formula for the ARL given by
E(N)=eta*inv(I-Q)*one;
identity = I(NN);

* Calculating the 5th, 25th, 50th, 75th and 95th percentiles;
p5p=0.05;
p25p=0.25;
p50p=0.5;
p75p=0.75;
p95p=0.95;
```

```
pmf=j(z,1,1);
cdf=j(z,1,1);
cdf_5th_p=j(z,1,1);
cdf_25th_p=j(z,1,1);
cdf_50th_p=j(z,1,1);
cdf_75th_p=j(z,1,1);
cdf_95th_p=j(z,1,1);

do i = 1 to z;
      pmf[i,1] = eta * (Q**(i-1)) * (identity - Q) * one;
      cdf[i,1]=sum(pmf[1:i,1]);
end;

index=j(z,1,1);

do i = 2 to z;
index[i,]=index[i-1,]+1;
end;

* Calculating the 5th percentile;
do i = 1 to z;
      cdf_5th_p[i,]=cdf[i,];
      if cdf_5th_p[i,]>=p5p then cdf_5th_p[i,]=999;
end;

cdf_5th_p=cdf_5th_p[loc(cdf_5th_p<999)];
if cdf_5th_p[1,]=999 then percentile_p5p=1;
else percentile_p5p=nrow(cdf_5th_p)+1;

* Calculating the 25th percentile;
do i = 1 to z;
      cdf_25th_p[i,]=cdf[i,];
      if cdf_25th_p[i,]>=p25p then cdf_25th_p[i,]=999;
end;

cdf_25th_p=cdf_25th_p[loc(cdf_25th_p<999)];
if cdf_25th_p[1,]=999 then percentile_p25p=1;
else percentile_p25p=nrow(cdf_25th_p)+1;

* Calculating the 50th percentile;
do i = 1 to z;
      cdf_50th_p[i,]=cdf[i,];
      if cdf_50th_p[i,]>=p50p then cdf_50th_p[i,]=999;
end;

cdf_50th_p=cdf_50th_p[loc(cdf_50th_p<999)];
if cdf_50th_p[1,]=999 then percentile_p50p=1;
else percentile_p50p=nrow(cdf_50th_p)+1;

* Calculating the 75th percentile;
do i = 1 to z;
      cdf_75th_p[i,]=cdf[i,];
      if cdf_75th_p[i,]>=p75p then cdf_75th_p[i,]=999;
end;

cdf_75th_p=cdf_75th_p[loc(cdf_75th_p<999)];
```

```
if cdf_75th_p[1,]=999 then percentile_p75p=1;
else percentile_p75p=nrow(cdf_75th_p)+1;

* Calculating the 95th percentile;
do i = 1 to z;
      cdf_95th_p[i,]=cdf[i,];
      if cdf_95th_p[i,]>=p95p then cdf_95th_p[i,]=999;
end;

cdf_95th_p=cdf_95th_p[loc(cdf_95th_p<999)];
if cdf_95th_p[1,]=999 then percentile_p95p=1;
else percentile_p95p=nrow(cdf_95th_p)+1;

* Calculating the average run length (ARL);
ARL = eta*ginv(identity-Q)*one;

* Calculating the second moment;
N2 = eta * (identity + Q) * (ginv((identity-Q)**2)) * one;

* Calculating the standard deviation;
SDRL = sqrt (N2 - ((ARL)**2) );

* Printing the output;
print_cdf=index||cdf;
print_pmf=index||pmf;

test = inv(identity - Q);

print       test,
            NN [label = 'Number of intervals between LCL and UCL'],
            n [label = 'Sample size'],
            L [label = 'L: EWMA parameter'],
            lambda [label = 'lambda: EWMA parameter'],
            UCL [label='Upper control limit'],
            LCL [label='Lower control limit'],
            delta,
            Q [format=fract.],
            ARL [label = 'Average run length' format=.2]
            SDRL [label = 'Standard deviation of the run length' format=.2],
            percentile_p5p [label='Fifth percentile'],
            percentile_p25p [label='25th percentile'],
            percentile_p50p [label='50th percentile'],
            percentile_p75p [label='75th percentile'],
            percentile_p95p [label='95th percentile'];
```

SAS program 9:

This program calculates the values for the CUSUM Mann-Whitney chart shown in Table 6.9.

```
proc iml;

* The data used by Sullivan and Woodall (1996) and Zhou, Zou and Wang (2007);
x_obs={
-0.69, 0.56, -0.96, -0.11, -0.25, 0.45, -0.26, 0.68, 0.22, -2.10, 0.65, -1.49,
-2.49, -1.11, 0.23, 2.16, 1.95, 1.54, 0.67, 1.09, 1.37, 0.69, 2.26, 1.86,
0.62, -1.04, 2.30, 0.07, 1.49, 0.52};

* Reference value;
k = 2;

* Calculating the Mann-Whitney (MW) values;
MW=j(nrow(x_obs)-1,1,.);

do r = 1 to nrow(x_obs-1);
      keep=0;
            do i = 1 to r;
                  count = j(nrow(x_obs),1,.);
                        do j = r+1 to nrow(x_obs);
                              if x_obs[i,]>x_obs[j,] then count[j,]=1;
                              t_sum=count[+,];
                        end;
                  keep = keep // t_sum;
            end;
      MW[r,] = keep[+,];
end;

t = j(nrow(MW),1,.);

do l = 1 to nrow(MW);
      t[l,]=l;
end;

MW = t||MW;

* Calculating the expected value of MWt, i.e. E(MWt);
exp = j(nrow(MW),1,.);

do l = 1 to nrow(MW);
      exp[l,]=(MW[l,1]*((nrow(x_obs))-MW[l,1]))/2;
end;

* Calculating the standard deviation of MWt, i.e. stdev(MWt);
stdev = j(nrow(MW),1,.);

do l = 1 to nrow(MW);
      stdev[l,]=sqrt((MW[l,1]*((nrow(x_obs))-MW[l,1])*(nrow(x_obs)+1))/12);
end;

* Calculating the standardized MWt values, i.e. SMWt;
SMW = j(nrow(MW),1,.);
```

```
do l = 1 to nrow(MW);
      SMW[l,]=(MW[l,2]-exp[l,])/stdev[l,];
end;

* Starting values for CUSUM;
S_plus = j (nrow(MW),1,.);
S_plus[1,]=0-SMW[1,]-k;
if S_plus[1,]<0 then S_plus[1,]=0;

S_minus = j (nrow(MW),1,.);
S_minus[1,]=0+SMW[1,]-k;
if S_minus[1,]<0 then S_minus[1,]=0;

* Calculating the CUSUM statistics;
do l = 2 to nrow(S_plus);
      S_plus[l,]=S_plus[l-1,]+SMW[l,]-k;
      if S_plus[l,]<0 then S_plus[l,]=0;
end;

do l = 2 to nrow(S_minus);
      S_minus[l,]=S_minus[l-1,]+SMW[l,]+k;
      if S_minus[l,]>0 then S_minus[l,]=0;
end;

MW = MW[,2];

print       x_obs [label='Xi-values' format=.2],
            MW [label='Mann-Whitney statistics'],
            exp [label='Expected values of MW-statistics'],
            stdev [label='Standard deviation values of the MW-statistics'
format=.3],
            SMW [label='Standarddized values for the MW-statistics'
format=.3],
            S_plus [label='Upper CUSUM statistics' format=.3],
            S_minus [label='Lower CUSUM statistics' format=.3];
```

SAS program 10:

This program calculates the charting statistics for the Shewhart-type signed-rank-like chart.

```
proc iml;

* Number of samples;
nn=15;

* Sample size;
n=5;

wsrmatrix = j(nn,n,0);
sgnmatrix = j(nn,n,0);
rank_abs_diff = j(nn,n,0);
final_rank_abs_diff=j(nn,n,0);

* Median of the reference sample;
tv=74.001;

tvmatrix = j(nn,n,tv);

* A matrix containing the Montgomery (2001) Table 5.2 piston ring data;
obs = {
74.012 74.015 74.030 73.986 74.000,
73.995 74.010 73.990 74.015 74.001,
73.987 73.999 73.985 74.000 73.990,
74.008 74.010 74.003 73.991 74.006,
74.003 74.000 74.001 73.986 73.997,
73.994 74.003 74.015 74.020 74.004,
74.008 74.002 74.018 73.995 74.005,
74.001 74.004 73.990 73.996 73.998,
74.015 74.000 74.016 74.025 74.000,
74.030 74.005 74.000 74.016 74.012,
74.001 73.990 73.995 74.010 74.024,
74.015 74.020 74.024 74.005 74.019,
74.035 74.010 74.012 74.015 74.026,
74.017 74.013 74.036 74.025 74.026,
74.010 74.005 74.029 74.000 74.020};

* Calculating the SRLi statistics;
diff=obs-tvmatrix;

do k = 1 to nn;
      do l = 1 to n;
      if diff[k,l]>0 then sgnmatrix[k,l]=1;
      else if diff[k,l]<0 then sgnmatrix[k,l]=-1;
      else if diff[k,l]=0 then sgnmatrix[k,l]=0;
      end;
end;

abs_diff=abs(diff);

do i = 1 to nn;
      rank_abs_diff[i,]=rank(abs_diff[i,]);
```

```
end;

do i = 1 to nn;
      do j = 1 to n;
            do k = 1 to n;
                  if abs_diff[i,j]=abs_diff[i,k] then

      final_rank_abs_diff[i,j]=(rank_abs_diff[i,j]+rank_abs_diff[i,k])/2;
            end;
      end;
end;

do i = 1 to nn;
      do j = 1 to n;
            do k = 1 to n;
                  if abs_diff[i,j]=abs_diff[i,k] then
                  final_rank_abs_diff[i,k]=final_rank_abs_diff[i,j];
            end;
      end;
end;


do k = 1 to nn;
      do l = 1 to n;
            wsrmatrix[k,l]=sgnmatrix[k,l]*final_rank_abs_diff[k,l];
      end;
end;

SRi = wsrmatrix[,+];

print SRi [label='Signed-rank-like statistics'];
```

Mathematica program 1:

Chakraborti and Van de Wiel (2003) wrote a Mathematica program which deals with the computation of the upper and lower control limits of the Mann-Whitney control chart for either a specified in-control *ARL* or a specified ρ^{th} percentile of conditional *ARL* (that is: that level of which one wants to be $100(1-\rho)\%$ sure that it is exceeded for his/her specific reference sample). In addition, it contains procedures for approximation of the ARL_0 and percentiles when control limits are given as well as procedures for computations under out-of-control situations. This Mathematica program can be reached using the website www.win.tue.nl/~markvdw. This program has user friendly parameters which I changed to suit my examples in Section 6.1.

References

Albers, W. and Kallenberg, W.C.M. (2004). "Empirical nonparametric control charts: estimation effects and corrections." *Journal of Applied Statistics,* **31**, 345-360.

Amin, R.W. (1987). "Variable sampling interval control charts." Unpublished Ph.D dissertation, Virginia Polytechnic Institute and State University, Department of Statistics.

Amin, R.W. and Hemasinha, R. (1993). "The switching behavior of $\overline{X}$ charts with variable sampling intervals." *Communications in Statistics: Theory and Methods,* **22** (7), 2081-2102.

Amin, R.W., Reynolds Jr. M.R. and Bakir, S.T. (1995). "Nonparametric quality control charts based on the sign statistic." *Communications in Statistics: Theory and Methods,* **24** (6), 1597-1623.

Amin, R.W. and Widmaier, O. (1999). "Sign control charts with variable sampling intervals." *Communications in Statistics: Theory and Methods,* **28**, 1961-1985.

Bain, L.J. and Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics, Second Edition,* Duxbury Press.

Bakir, S.T. (2003). "A quality control chart based on signed-ranks." *Joint Statistical Meeting – Section on Quality and Productivity,* 423-429.

Bakir, S.T. (2004). "A distribution-free Shewhart quality control chart based on signed-ranks." *Quality Engineering,* **16**, 613-623.

Bakir, S.T. (2006). "Distribution-free quality control charts based on signed-rank-like statistics." *Communications in Statistics: Theory and Methods*, **35**, 743-757.

Bakir, S.T. and Reynolds, Jr. M.R. (1979). "A nonparametric procedure for proses control based on within-group ranking." *Technometrics*, **21** (2), 175-183.

Barnard, G.A. (1959). "Control charts and stochastic processes." *Journal of the Royal Statistical Society, B,* **21**, 239-271.

Bartlett, M.S. (1953). "Recurrence and first passage times." *Proceedings of the Cambridge Philosophical Society,* **49**, 263-275.

Borror, C.M., Montgomery, D.C. and Runger, G.C. (1999). "Robustness of EWMA control charts to non-normality." *Journal of Quality Technology*, **31** (3), 309-316.

Brook, D. and Evans, D.A. (1972). "An approach to the probability distribution of CUSUM run length." *Biometrika*, **59** (3), 539-549.

Chakraborti, S. (2000). "Run length, average run length and false alarm rate of Shewhart $\overline{X}$ chart: Exact derivations by conditioning." *Communications in Statistics: Simulation and Computation,* **29** (1), 61-81.

Chakraborti, S. (2006). "Parameter estimation and design considerations in prospective applications of the $\overline{X}$ chart." *Journal of Applied Statistics,* **33**, 439-459.

Chakraborti, S. (2007). "Run length distribution and percentiles: The Shewhart $\overline{X}$ chart with unknown parameters." *Quality Engineering,* **19**, 119-127.

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

Chakraborti, S., Eryilmaz, S. and Human, S.W. (2006). "A Phase II Shewhart-type nonparametric control chart based on runs." (Submitted).

Chakraborti, S. and Van der Laan, P. (1996). "Precedence tests and confidence bounds for complete data: an overview and some results." *Journal of the Royal Statistical Society Series D: The Statistician,* **45** (3), 351-369.

Chakraborti, S. and Van der Laan, P. (1997). "An overview of precedence tests for censored data." *Biometrical Journal,* **39** (1), 99-116.

Chakraborti, S., Van der laan, P. and Bakir, S.T. (2001). "Nonparametric control charts: An overview and some results." *Journal of Quality Technology* **33** (3), 304-315.

Chakraborti, S., Van der Laan, P. and Van de Wiel, M.A. (2004). "A class of distribution-free control charts." *Journal of the Royal Statistical Society. Series C*: *Applied Statistics*, **53**, 443-462.

Chakraborti, S. and Van de Wiel, M.A. (2003). "A nonparametric control chart based on the Mann-Whitney statistic." SPOR Report 2003-24, Technical University of Eindhoven, Eindhoven.

Champ, C.W. and Chou, S.P. (2003). "Comparison of standard and individual limits Phase I Shewhart $\overline{X}$, *R* and *S* charts." *Quality and Reliability Engineering International*, **19**, 161-170.

Champ, C.W. and Jones, L.A. (2004). "Designing Phase I $\overline{X}$ chart with small sample sizes." *Quality and Reliability Engineering International,* **20**, 497-510.

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

Champ, C.W. and Woodall, W.H. (1997). "Signal probabilities of runs rules supplementing a Shewhart control chart." *Communications in Statistics - Simulation and Computation,* **26** (4), 1347-1360.

Chou, S.P. and Champ, C.W. (1995). "A comparison of two Phase I control charts." *Proceedings of the Quality and Productivity Section of the American Statistical Association*, 31-35.

Crowder, S.V. (1987). "Simple methods for studying run length distributions of exponentially weighted moving average control charts." *Technometrics,* **29**, 401-407.

Daniels, H.E. (1954). "Saddlepoint approximations in statistics." *Annals of Mathematical Statistics,* **25**, 631-650.

Daniels, H.E. (1987). "Tail probability approximations." *International Statistical Review,* **55** (1), 37-48.

Ewan, W. D. (1963). "When and how to use CUSUM charts." *Technometrics*, **5** (1), 1-22.

Ewan, W.D. and Kemp, K.W. (1960). "Sampling inspection of continuous processes with no autocorrelation between successive results." *Biometrika*, **47**, 363-380.

Fix, E. and Hodges, J.L. (1955). "Significance probabilities of the Wilcoxon test." *Annals of Statistics,* **26**, 301-312.

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

Fu, J.C., Spiring, F.A. and Xie, H. (2002). "On the average run lengths of quality control schemes using a Markov chain approach." *Statistics and Probability Letters,* **56**, 369-380.

Gan, F.F. (1993). "An optimal design of CUSUM control charts for binomial counts." *Journal of Applied Statistics,* **20**, 445-460.

Gibbons, J.D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference,* 4th Edition, Revised and Expanded, Marcel Dekker, New York.

Goldsmith, R.L. and Whitfield, H. (1961). "Average run lengths in cumulative chart quality control schemes." *Technometrics,* **3** (1), 11-20.

Hawkins, D.M. (1977). "Testing a sequence of observations for a shift in location." *Journal of the American Statistical Association*, **72**, 180-186.

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

Hawkins, D.M., Qiu, P. and Kang, C.W. (2003). "The change-point model for statistical process control." *Journal of Quality Technology*, **35**, 355-366.

Hawkins, D.M. and Zamba, K.D. (2005). "Statistical process control for shifts in mean or variance using a change-point formulation." *Technometrics*, **47**, 164-173.

Hollander, M. and Wolfe, D.A. (1973). *Nonparametric statistics.* New York: John Wiley.

Human, S.W., Chakraborti, S. and Smit, C.F. (2007). "Design of S^2, S and R control charts for Phase I application." Submitted.

Human, S.W., Chakraborti, S. and Smit, C.F. (2008). "Nonparametric Shewhart-type sign control charts based on runs." Submitted.

Janacek, G.J. and Meikle, S.E. (1997). "Control charts based on medians." *The Statistician*, **46** (1), 19-31.

Jensen, J.L. (1995). *Saddlepoint Approximations,* Oxford Statistical Science Series, Clarendon Press, Oxford.

Johnson, N.L. (1961). "Simple theoretical approach to cumulative sum control charts." *Journal of the American Statistical Association,* **56**, 835-840.

Jones, L.A. (2002). "The statistical design of EWMA control charts with estimated parameters." *Journal of Quality Technology,* **34** (3), 277-288.

Jones, L.A. and Champ, C.W. (2002). "Phase I control charts for times between events." *Quality and Reliability Engineering International,* **18**, 479-488.

Kemp, K.W. (1971). "Formal expressions which can be applied to CUSUM charts." *Journal of the Royal Statistical Society, Series B: Methodological,* **33**, 331-360.

Khoo, M.B.C. (2004). "Design of runs rules schemes." *Quality Engineering* **16** (1), 27-43.

King, E.P. (1954). "Probability limits for the average chart when process standards are unspecified." *Industrial Quality Control*, **10**, 62-64.

Klein, M. (2000). "Two alternatives to the Shewhart $\overline{X}$ control chart." *Journal of Quality Technology*, **32** (4), 427-431.

Koning, A.J. (2006). "Model-based control charts in Phase I statistical process control." *Statistica Neerlandica,* **60** (3), 327-338.

Koning, A.J. and Does, R.J.M.M. (2000). "CUSUM charts for preliminary analysis of individual observations." *Journal of Quality Technology*, **32**, 122-132.

Lehmann, E.L. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco: McGraw-Hill.

Lucas, J.M. (1985). "Counted data CUSUMS." *Technometrics,* **27** (2), 129-144.

Lucas, J.M. and Crosier, R.B. (1982). "Fast initial response for CUSUM quality-control schemes: Give your CUSUM a head start." *Technometrics*, **24** (3), 199-205.

Lucas, J.M. and Saccucci, M.S. (1987). "Exponentially weighted moving average control schemes: Properties and enhancements." *Faculty Working Series Paper 87-5*, Drexel University, Department of Quantitative Methods.

Lucas, J.M. and Saccucci, M.S. (1990). "Exponentially weighted moving average control schemes: Properties and enhancements." *Technometrics,* **32** (1), 1-12.

Lugannani, R. and Rice, S. (1980). "Saddlepoint approximation for the distribution of the sum of independent random variables." *Advances in Applied Probability,* **12**, 475-490.

Mathisen, H.C. (1943). "A method of testing the hypothesis that two samples are from the same population." *Annals of Mathematical Statistics,* **14**, 188-194.

Montgomery, D.C. (2001). *Introduction to Statistical Quality Control,* 4th Edition, John Wiley & Sons, New York.

Montgomery, D.C. (2005). *Introduction to Statistical Quality Control 5th ed.* New York: John Wiley.

Montgomery, D.C., Gardiner, J.S. and Pizzano, B.A. (1987). "Statistical process control methods for detecting small process shifts." *Frontiers in Statistical Quality Control,* **3,** 163-178.

Moses, L.E. (1963). "Rank tests of dispersion." *The Annals of Mathematical Statistics*, **34** (3), 973-983.

Nelson, L.S. (1984). "The Shewhart control chart - tests for special causes." *Journal of Quality Technology*, **16**, 237-239.

Page, E.S. (1954). "Continuous inspection schemes." *Biometrika*, **41**, 100-115 .

Page, E.S. (1955). "Control charts with warning lines." *Biometrika,* **42**, 243-257.

Page, E. S. (1961). "Cumulative sum charts." *Technometrics*, **3** (1), 1-9.

Page, E. S. (1962). "A modified control chart with warning lines." *Biometrika*, **49**, 171-176.

Park, C. and Reynolds, Jr. M.R. (1987). "Nonparametric procedures for monitoring a location parameter based on linear placement statistics." *Sequential Analysis*, **6**, 303-323.

Pettitt, A.N. (1979). "A nonparametric approach to the change-point problem." *Applied Statistics*, **28**, 126-135.

Radson, D. and Boyd, A.H. (2005). "Graphical representation of run length distributions." *Quality Engineering*, **17**, 301-308.

Randles, R.H. and Wolfe, D.A. (1979). *Introduction to the Theory of Nonparametric Statistics,* John Wiley and Sons, New York.

Reid, N. (1988). "Saddlepoint methods and statistical inference (with discussion)." *Statistical Science*, **3**, 213-238.

Rendtel, U. (1990). "CUSUM schemes with variable sampling intervals and sampling sizes." *Statistical Papers*, **31**, 103-118.

Reynolds, M.R., Amin, R.W. and Arnold, J.C. (1990). "CUSUM charts with variable sampling intervals." *Technometrics*, **32** (4), 371-384.

Roberts, S.W. (1958). "Properties of control chart zone tests." *The Bell System Technical Journal* **37**, 83-114.

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

Roberts, S.W. (1966). "A comparison of some control chart procedures." *Technometrics,* **8** (3)**,** 411-430.

Robinson, P.B. and Ho, T.Y. (1978). "Average run lengths of geometric moving average charts by numerical methods." *Technometrics*, **20**, 85-93.

Saccucci, M.S., Amin, R.W. and Lucas, J.M. (1992). "Exponentially weighted moving average control schemes with variable sampling intervals." *Communications in Statistics: Simulation and Computation,* **21** (3), 627-657.

Sheu, S. and Yang, L. (2006). "The generally weighted moving average median control chart." *Quality Technology and Quantitative Management,* **3** (4), 455-471.

Shewhart, W.A. (1926). "Quality control charts." *Bell Systems Technical Journal*, 593-603.

Shewhart, W.A. (1931). *Economic Control of Quality of Manufactured Product,* New York: Van Nostrand.

Shewhart, W.A. (1939). "Economic control of quality of manufactured product." *Princeton: Van Nostrand Reinhold.*

Shewhart, W.A. (1941). "Contributions of statistics to the science of engineering." *Fluid Mechanics and Statistical Methods in Engineering,* University of Pennsylvania Press, Philadelphia, 97-124.

Shmueli, G. and Cohen, A. (2003). "Run-length distribution for control charts with runs and scans rules." *Communications in Statistics: Theory and Methods,* **32** (2), 475-495.

Sullivan, J.H. and Woodall, W.H. (1996). "A control chart for preliminary analysis of individual observations." *Journal of Quality Technology*, **28** (3), 265-278.

Van de Wiel, M.A. (2007). Personal communication.

Van der Laan, P. and Chakraborti, S. (1999). "Best precedence tests against Lehmann alternatives." *Bulletin International Statistical Institute*, **58** (3), 395-396.

Van Dobben de Bruyn, C.S. (1968). *Cumulative Sum Tests, New York: Hafner.*

Weindling, J.I., Littauer, S.B. and Oliveira, J.T. (1979). "Mean action time of the $\overline{X}$ control chart with warning limits." *Journal of Quality Technology*, **2** (2), 79-85.

Western Electronic Company. (1956). *Statistical Quality Control Handbook,* Western Electric Corporation, Indianaplois.

White, C. H., Keats, J. B. and Stanley, J. (1997). "Poisson CUSUM versus c-chart for defect data." *Quality Engineering*, **9**, 673-679.

Woodall, W.H. (1983). "The distribution of the run length of one-sided CUSUM procedures for continuous random variables." *Technometrics*, **25**, 295-301

Woodall, W. and Adams, B. (1993). "The statistical design of CUSUM charts." *Quality Engineering*, **5**, 559-570.

Woodall, W.H. and Montgomery, D.C. (1999). "Research issues and ideas in statistical process control." *Journal of Quality Technology*, **31** (4), 376-386.

Woodall, W.H. (2000). "Controversies and contradictions in statistical process control." *Journal of Quality Technology*, **32** (4), 341-350.

Wu, C., Zhao, Y. and Wang, Z. (2002). "The median absolute deviations and their applications to Shewhart $\overline{X}$ control charts." *Communications in Statistics: Simulation and Computation,* **31** (3), 425-442.

Zhou, C., Zou, C. and Wang, Z. "Nonparametric control chart for preliminary analysis of individual observations." Preprint, http://202.113.29.3/keyan/pre/preprint06/06-17.pdf, December 2007.